OM of: US-09-439-311-2 to: Issued_Patents_NA:* out_format : pfs

Date: Apr 17, 2002 3:10 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09439311/runat_16042002_134010_11703/app_query.fasta_1.395
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=60 -USER=US09439311_@CGN1_1_77
-NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-439-311-2
Query length: 333
Database: Issued_Patents_NA:*
Database sequences: 351203
Database length: 113238999
Search time (sec): 85.700000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	!	Documentation ..
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-358-972-102	-	50.00	116.12	80.95	30	!	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-358-972-103	+	50.00	116.12	80.95	30	!	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-406-147-32	-	50.00	116.12	80.95	30	!	S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-406-147-34	+	50.00	116.12	80.95	30	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-130-663-22	+	43.00	96.03	1.1e+03	51	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-081-180-29	+	43.00	96.03	1.1e+03	51	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-040-786-29	+	43.00	96.03	1.1e+03	51	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-432-335-22	+	43.00	96.03	1.1e+03	51	!	S
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-219-012-80	+	40.00	88.19	2.9e+03	60	!	S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-687-421-268	+	40.00	88.19	2.9e+03	60	!	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-482-882-97	-	39.00	89.11	2.6e+03	47	!	S
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-389-97	-	39.00	89.11	2.6e+03	47	!	S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-487-113D-97	-	39.00	89.11	2.6e+03	47	!	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-473-503-97	-	39.00	89.11	2.6e+03	47	!	S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-483-932-97	-	39.00	89.11	2.6e+03	47	!	S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-720-420A-97	-	39.00	89.11	2.6e+03	47	!	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-714-017-97	-	39.00	89.11	2.6e+03	47	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-475-680-97	-	39.00	89.11	2.6e+03	47	!	S
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-537-357-46	+	38.00	87.66	3.1e+03	45	!	S
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-353-657-3	-	38.00	85.71	4.0e+03	53	!	Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-709-982-3	-	38.00	85.71	4.0e+03	53	!	Se
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-982-866-3	-	38.00	85.71	4.0e+03	53	!	Se
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-14181-3	-	38.00	85.71	4.0e+03	53	!	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-587-209-6	-	37.00	93.14	1.5e+03	24	!	Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-689-236-6	-	37.00	93.14	1.5e+03	24	!	Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-689-235-6	-	37.00	93.14	1.5e+03	24	!	Se
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-692-725-6	-	37.00	93.14	1.5e+03	24	!	Se
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-692-726-6	-	37.00	93.14	1.5e+03	24	!	Se
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-813-507-82	-	37.00	86.79	3.5e+03	41	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-109-063-28	+	37.00	84.67	4.6e+03	49	!	S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-975-699-4	-	37.00	82.27	6.2e+03	60	!	Se
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-972-089-4	-	37.00	82.27	6.2e+03	60	!	Se
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-766-354A-15	-	37.00	82.27	6.2e+03	60	!	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-133-717-5	-	37.00	82.27	6.2e+03	60	!	Se
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-158-863C-5	-	37.00	82.27	6.2e+03	60	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-675-566-100	+	36.50	81.89	6.5e+03	57	!	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-675-566-101	+	36.50	81.89	6.5e+03	57	!	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-564-955-54	+	36.00	81.77	6.6e+03	53	!	S
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-537-874-54	+	36.00	81.77	6.6e+03	53	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-621-859-54	+	36.00	81.77	6.6e+03	53	!	S
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-100-856A-54	+	36.00	81.77	6.6e+03	53	!	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-075-511-54	+	36.00	81.77	6.6e+03	53	!	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-099-015-54	+	36.00	81.77	6.6e+03	53		
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-232-863-54	+	36.00	81.77	6.6e+03	53		
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-133-508A-54	+	36.00	81.77	6.6e+03	53		

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-358-972-102

seq_documentation_block:
; Sequence 102, Application US/09358972
; Patent No. 6235480
; GENERAL INFORMATION:
; APPLICANT: Shultz, John W
; APPLICANT: Lewis, Martin K.
; APPLICANT: Lieppe, Donna
; APPLICANT: Mandrekar, Michelle
; APPLICANT: Kephart, Daniel
; APPLICANT: Rhodes, Richard B.
; APPLICANT: Andrews, Christine A.
; APPLICANT: Hartnett, James R.
; APPLICANT: Gu, Trent
; APPLICANT: Olson, Ryan J.
; APPLICANT: Wood, Keith W.
; APPLICANT: Welch, Roy
; TITLE OF INVENTION: Nucleic Acid Detection
; FILE REFERENCE: Pro-103 6868/75528
; CURRENT APPLICATION NUMBER: US/09/358,972
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: 09/252,436
; EARLIER FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: 09/042,287
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: probe to Campylobacter jejuni
US-09-358-972-102

alignment_scores:
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-439-311-2 x US-09-358-972-102/rev ..

Align seg 1/1 to reverse of: US-09-358-972-102 from: 1 to: 30

```
      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
      30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1
```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-358-972-103

seq_documentation_block:
; Sequence 103, Application US/09358972
; Patent No. 6235480
; GENERAL INFORMATION:
; APPLICANT: Shultz, John W
; APPLICANT: Lewis, Martin K.
; APPLICANT: Lieppe, Donna
; APPLICANT: Mandrekar, Michelle
; APPLICANT: Kephart, Daniel
; APPLICANT: Rhodes, Richard B.
; APPLICANT: Andrews, Christine A.
; APPLICANT: Hartnett, James R.
; APPLICANT: Gu, Trent
; APPLICANT: Olson, Ryan J.
; APPLICANT: Wood, Keith W.

```
;  APPLICANT: Welch, Roy
;  TITLE OF INVENTION: Nucleic Acid Detection
;  FILE REFERENCE: Pro-103 6868/75528
;  CURRENT APPLICATION NUMBER: US/09/358,972
;  CURRENT FILING DATE: 1999-07-22
;  EARLIER APPLICATION NUMBER: 09/252,436
;  EARLIER FILING DATE: 1999-02-18
;  EARLIER APPLICATION NUMBER: 09/042,287
;  EARLIER FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 290
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Campylobacter jejuni
;   FEATURE:
;   OTHER INFORMATION: probe to Campylobacter jejuni
US-09-358-972-103


alignment_scores:
            Quality:   50.00              Length:     10
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-439-311-2 x US-09-358-972-103    ..

 Align seg 1/1  to: US-09-358-972-103  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
       1 CAAGATGGACAAAGTTTAAAAACAAGAACT 30

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-406-147-32

seq_documentation_block:
; Sequence 32, Application US/09406147
; Patent No. 6270974
; GENERAL INFORMATION:
;  APPLICANT: Shultz, John W
;  APPLICANT: Lewis, Martin K
;  APPLICANT: Leippe, Donna
;  APPLICANT: Mandrekar, Michelle
;  APPLICANT: Kephart, Daniel
;  APPLICANT: Rhodes, Richard B
;  APPLICANT: Andrews, Christine A
;  APPLICANT: Hartnett, James R
;  APPLICANT: Gu, Trent
;  APPLICANT: Wood, Keith V
;  APPLICANT: Welch, Roy
;  TITLE OF INVENTION: EXOGENOUS NUCLEIC ACID DETECTION
;  FILE REFERENCE: EXOGENOUS NUCLEIC ACID DETECTION
;  CURRENT APPLICATION NUMBER: US/09/406,147
;  CURRENT FILING DATE: 1999-09-27
;  EARLIER APPLICATION NUMBER: 09/252,436
;  EARLIER FILING DATE: 1999-02-18
;  EARLIER APPLICATION NUMBER: 09/042,287
;  EARLIER FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Campylobacter jejuni
US-09-406-147-32


alignment_scores:
            Quality:   50.00              Length:     10
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-439-311-2 x US-09-406-147-32/rev    ..

 Align seg 1/1  to reverse of: US-09-406-147-32  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
       30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-406-147-34

seq_documentation_block:
; Sequence 34, Application US/09406147
; Patent No. 6270974
; GENERAL INFORMATION:
;  APPLICANT: Shultz, John W
;  APPLICANT: Lewis, Martin K
;  APPLICANT: Leippe, Donna
;  APPLICANT: Mandrekar, Michelle
;  APPLICANT: Kephart, Daniel
;  APPLICANT: Rhodes, Richard B
;  APPLICANT: Andrews, Christine A
;  APPLICANT: Hartnett, James R
;  APPLICANT: Gu, Trent
;  APPLICANT: Wood, Keith V
;  APPLICANT: Welch, Roy
;  TITLE OF INVENTION: EXOGENOUS NUCLEIC ACID DETECTION
;  FILE REFERENCE: EXOGENOUS NUCLEIC ACID DETECTION
;  CURRENT APPLICATION NUMBER: US/09/406,147
;  CURRENT FILING DATE: 1999-09-27
;  EARLIER APPLICATION NUMBER: 09/252,436
;  EARLIER FILING DATE: 1999-02-18
;  EARLIER APPLICATION NUMBER: 09/042,287
;  EARLIER FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Campylobacter jejuni
US-09-406-147-34


alignment_scores:
            Quality:   50.00              Length:     10
              Ratio:  5.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-439-311-2 x US-09-406-147-34    ..

 Align seg 1/1  to: US-09-406-147-34  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
       1 CAAGATGGACAAAGTTTAAAAACAAGAACT 30

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-130-663-22

seq_documentation_block:
; Sequence 22, Application US/09130663A
; Patent No. 6020163
; GENERAL INFORMATION:
;  APPLICANT: Conklin, Darrell C.
;  TITLE OF INVENTION: LIPOCALIN HOMOLOG
;  FILE REFERENCE: 97-24
;  CURRENT APPLICATION NUMBER: US/09/130,663A
;  CURRENT FILING DATE: 1998-08-05
;  EARLIER APPLICATION NUMBER: 60/054,867
;  EARLIER FILING DATE: 1997-08-06
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
```

```
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide primer: ZC13735.
US-09-130-663-22


alignment_scores:
            Quality:   43.00              Length:     16
              Ratio:  3.583                Gaps:      0
 Percent Similarity:  75.000   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x US-09-130-663-22    ..

 Align seg 1/1  to: US-09-130-663-22  from: 1  to: 51

     253 GlyValValIleGlyLysValAspTyrSerAspGlyAspGluAsnGly 268
         ||||||   :::   |||   |||||||:::|||   |||:::::|||
       1 GGTGTAAGCTTGGACAAGAGAGATTACAAGGACGATGATGACAAGGGT 48

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-081-180-29

seq_documentation_block:
; Sequence 29, Application US/09081180
; Patent No. 6022847
;  GENERAL INFORMATION:
;    APPLICANT:  Sheppard, Paul O.
;    APPLICANT:  .
;    TITLE OF INVENTION:  SECRETED SALIVARY ZSIG32
;    TITLE OF INVENTION:  POLYPEPTIDES
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics
;      STREET:  1201 Eastlake Ave. E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/081,180
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/041,263
;      FILING DATE:  March 19, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lingenfelter, Susan E
;      REGISTRATION NUMBER:  41,156
;      REFERENCE/DOCKET NUMBER:  97-17C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6675
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  ZC13735
US-09-081-180-29


alignment_scores:
            Quality:   43.00              Length:     16
              Ratio:  3.583                Gaps:      0
 Percent Similarity:  75.000   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x US-09-081-180-29    ..

 Align seg 1/1  to: US-09-081-180-29  from: 1  to: 51

     253 GlyValValIleGlyLysValAspTyrSerAspGlyAspGluAsnGly 268
         ||||||   :::   |||   |||||||:::|||   |||:::::|||
       1 GGTGTAAGCTTGGACAAGAGAGATTACAAGGACGATGATGACAAGGGT 48

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-040-786-29

seq_documentation_block:
; Sequence 29, Application US/09040786
; Patent No. 6025197
;  GENERAL INFORMATION:
;    APPLICANT:  Sheppard, Paul O.
;    APPLICANT:  .
;    TITLE OF INVENTION:  SECRETED SALIVARY ZSIG32
;    TITLE OF INVENTION:  POLYPEPTIDES
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics
;      STREET:  1201 Eastlake Ave. E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/040,786
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/041,263
;      FILING DATE:  March 19, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lingenfelter, Susan E
;      REGISTRATION NUMBER:  41,156
;      REFERENCE/DOCKET NUMBER:  97-17
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6675
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  ZC13735
US-09-040-786-29


alignment_scores:
            Quality:   43.00              Length:     16
              Ratio:  3.583                Gaps:      0
 Percent Similarity:  75.000   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x US-09-040-786-29    ..

 Align seg 1/1  to: US-09-040-786-29  from: 1  to: 51
```

```
     253 GlyValValIleGlyLysValAspTyrSerAspGlyAspGluAsnGly 268
         ||||||   :::   |||   ||||||:::|||   |||:::::::|||
       1 GGTGTAAGCTTGGACAAGAGAGATTACAAGGACGATGATGACAAGGGT 48

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-432-335-22

seq_documentation_block:
; Sequence 22, Application US/09432335
; Patent No. 6143720
; GENERAL INFORMATION:
;  APPLICANT: Conklin, Darrell C.
;  TITLE OF INVENTION: LIPOCALIN HOMOLOG
;  FILE REFERENCE: 97-24
;  CURRENT APPLICATION NUMBER: US/09/432,335
;  CURRENT FILING DATE: 1999-11-02
;  EARLIER APPLICATION NUMBER: 09/130,663
;  EARLIER FILING DATE: 1998-08-06
;  EARLIER APPLICATION NUMBER: 60/054,867
;  EARLIER FILING DATE: 1997-08-06
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide primer: ZC13735
US-09-432-335-22


alignment_scores:
           Quality:   43.00            Length:     16
             Ratio:  3.583              Gaps:      0
 Percent Similarity:  75.000   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x US-09-432-335-22   ..

 Align seg 1/1  to: US-09-432-335-22  from: 1  to: 51

     253 GlyValValIleGlyLysValAspTyrSerAspGlyAspGluAsnGly 268
         ||||||   :::   |||   ||||||:::|||   |||:::::::|||
       1 GGTGTAAGCTTGGACAAGAGAGATTACAAGGACGATGATGACAAGGGT 48

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-219-012-80

seq_documentation_block:
; Sequence 80, Application US/08219012
; Patent No. 5543293
;  GENERAL INFORMATION:
;    APPLICANT:  Larry Gold
;    APPLICANT:  Diane Tasset
;    TITLE OF INVENTION:  Ligands of Thrombin
;    NUMBER OF SEQUENCES:  92
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Beaton & Swanson, P.C.
;      STREET:  4582 South Ulster Street Parkway, Suite #
;      STREET:  403
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80237
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 Kb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/219,012
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:  none
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 850-9900
;      TELEFAX:  (303) 850-9401
;  INFORMATION FOR SEQ ID NO:  80:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-219-012-80


alignment_scores:
           Quality:   40.00            Length:     19
             Ratio:  2.667              Gaps:      0
 Percent Similarity:  78.947   Percent Identity: 42.105

alignment_block:
 US-09-439-311-2 x US-08-219-012-80   ..

 Align seg 1/1  to: US-08-219-012-80  from: 1  to: 60

     204 ThrSerValGlyThrGlyLeuGlyAlaLeuAlaGluGluIleAsnArgAs 220
         |||:::   |||:::|||||   |||||||||:::   ::::::   ::
       4 ACCGCGGAGGGCGTAGGGTTGGAGGCGTTGGCCGATGTGGTAGGCACGGA 53

     220 nAlaAsp 222
         ::::|||
      54 CTCGGAT 60

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-687-421-268

seq_documentation_block:
; Sequence 268, Application US/08687421
; Patent No. 6177557
;  GENERAL INFORMATION:
;    APPLICANT:  Gold, Larry
;    APPLICANT:  Janjic, Nebojsa
;    APPLICANT:  Tasset, Diane
;    TITLE OF INVENTION:  HIGH-AFFINITY LIGANDS OF BASIC
;    TITLE OF INVENTION:  FIBROBLAST GROWTH FACTOR AND
;    TITLE OF INVENTION:  THROMBIN
;    NUMBER OF SEQUENCES:  445
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MB storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/687,421
;      FILING DATE:  08-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/195,005
;      FILING DATE:  10-FEBRUARY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:  22-APRIL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/219,012
;      FILING DATE:  28-MARCH-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/973,333
```

```
;       FILING DATE:  11-NOVEMBER-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/714,131
;       FILING DATE:  10-JUNE-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/536,428
;       FILING DATE:  11-JUNE-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Barry J. Swanson
;       REGISTRATION NUMBER:  33,215
;       REFERENCE/DOCKET NUMBER:  NEX07/PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (303) 793-3333
;       TELEFAX:  (303) 793-3433
;   INFORMATION FOR SEQ ID NO:  268:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  60 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-687-421-268


alignment_scores:
            Quality:   40.00              Length:     19
              Ratio:  2.667                Gaps:      0
 Percent Similarity:  78.947   Percent Identity: 42.105

alignment_block:
 US-09-439-311-2 x US-08-687-421-268   ..

 Align seg 1/1  to: US-08-687-421-268  from: 1  to: 60

     204 ThrSerValGlyThrGlyLeuGlyAlaLeuAlaGluGluIleAsnArgAs 220
         |||:::   |||:::||||||   ||||||||||:::   ::::::   ::
       4 ACCGCGGAGGGCGTAGGGTTGGAGGCGTTGGCCGATGTGGTAGGCACGGA 53

     220 nAlaAsp 222
         ::::|||
      54 CTCGGAT 60

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-482-882-97

seq_documentation_block:
; Sequence 97, Application US/08482882
; Patent No. 5773218
;  GENERAL INFORMATION:
;     APPLICANT:  Gallatin, W. Michael
;     APPLICANT:  Vazeux, Rosemay
;     TITLE OF INVENTION:  ICAM-Related Materials and Methods
;     NUMBER OF SEQUENCES:  116
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 S. Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60606
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/482,882
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/286,754
;       FILING DATE:
;       APPLICATION NUMBER:  US 08/102,852
;       FILING DATE:  05-AUG-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/009,266
;       FILING DATE:  22-JAN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/894,061
;       FILING DATE:  05-JUN-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/889,724
;       FILING DATE:  26-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/827,689
;       FILING DATE:  27-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5773218and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  32178
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  97:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-482-882-97


alignment_scores:
            Quality:   39.00              Length:     12
              Ratio:  3.900                Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x US-08-482-882-97/rev   ..

 Align seg 1/1  to reverse of: US-08-482-882-97  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::||||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-483-389-97

seq_documentation_block:
; Sequence 97, Application US/08483389
; Patent No. 5811517
;  GENERAL INFORMATION:
;     APPLICANT:  Gallatin, W. Michael
;     APPLICANT:  Vazeux, Rosemay
;     TITLE OF INVENTION:  ICAM-RELATED PROTEIN
;     NUMBER OF SEQUENCES:  118
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive/6300 Sears Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/483,389
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/102,852
;       FILING DATE:  05-AUG-1993
;     PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US 08/009,266
;       FILING DATE:  22-JAN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/894,061
;       FILING DATE:  05-JUN-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/889,724
;       FILING DATE:  26-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/827,689
;       FILING DATE:  27-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Suh, Young J.
;       REGISTRATION NUMBER:  P-41,337
;       REFERENCE/DOCKET NUMBER:  27866/32760
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  (312) 474-6600
;   INFORMATION FOR SEQ ID NO:  97:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-483-389-97

 

alignment_scores:
            Quality:   39.00              Length:     12
              Ratio:  3.900                Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x US-08-483-389-97/rev   ..

 Align seg 1/1  to reverse of: US-08-483-389-97  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||       |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-487-113D-97

seq_documentation_block:
; Sequence 97, Application US/08487113D
; Patent No. 5837822
;  GENERAL INFORMATION:
;     APPLICANT:  Gallatin, W. Michael
;     APPLICANT:  Vazeux, Rosemay
;     TITLE OF INVENTION:  ICAM-Related Materials and Methods
;     NUMBER OF SEQUENCES:  120
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 South Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/487,113D
;       FILING DATE:
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/286,754
;       FILING DATE:  05-AUG-1994
;     PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US 08/102,852
;       FILING DATE:  05-AUG-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/009,266
;       FILING DATE:  22-JAN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/894,061
;       FILING DATE:  05-JUN-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/889,724
;       FILING DATE:  26-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/827,689
;       FILING DATE:  27-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5837822and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  32744
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  97:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-487-113D-97


alignment_scores:
            Quality:   39.00              Length:     12
              Ratio:  3.900                Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x US-08-487-113D-97/rev   ..

 Align seg 1/1  to reverse of: US-08-487-113D-97  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||       |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-473-503-97

seq_documentation_block:
; Sequence 97, Application US/08473503
; Patent No. 5869262
;  GENERAL INFORMATION:
;     APPLICANT:  Gallatin, W. Michael
;     APPLICANT:  Vazeux, Rosemay
;     TITLE OF INVENTION:  ICAM-Related Materials and Methods
;     NUMBER OF SEQUENCES:  116
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 S. Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60606
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/473,503
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  08/286,754
;       FILING DATE:  05-AUG-1994
;       APPLICATION NUMBER:  US 08/102,852
;       FILING DATE:  05-AUG-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/009,266
;       FILING DATE:  22-JAN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/894,061
;       FILING DATE:  05-JUN-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/889,724
;       FILING DATE:  26-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/827,689
;       FILING DATE:  27-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5869262and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  32178
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  97:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-473-503-97


alignment_scores:
            Quality:   39.00              Length:     12
              Ratio:  3.900                Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x US-08-473-503-97/rev   ..

 Align seg 1/1  to reverse of: US-08-473-503-97  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-483-932-97

seq_documentation_block:
; Sequence 97, Application US/08483932
; Patent No. 5880268
;  GENERAL INFORMATION:
;     APPLICANT:  Gallatin, W. Michael
;     APPLICANT:  Vazeux, Rosemay
;     TITLE OF INVENTION:  ICAM-Related Materials and Methods
;     NUMBER OF SEQUENCES:  116
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 S. Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60606
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/483,932
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/286,754
;       FILING DATE:  05-AUG-1994
;       APPLICATION NUMBER:  US 08/102,852
;       FILING DATE:  05-AUG-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/009,266
;       FILING DATE:  22-JAN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/894,061
;       FILING DATE:  05-JUN-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/889,724
;       FILING DATE:  26-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/827,689
;       FILING DATE:  27-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5880268and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  32178
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  97:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-483-932-97


alignment_scores:
            Quality:   39.00              Length:     12
              Ratio:  3.900                Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x US-08-483-932-97/rev   ..

 Align seg 1/1  to reverse of: US-08-483-932-97  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9
```

```
OM of: US-09-439-311-2 to: GenEmbl:*   out_format : pfs

Date: Apr 17, 2002  3:08 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09439311/runat_16042002_134010_11688/app_query.fasta_1.395
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=60
 -USER=US09439311_@CGN1_1_5662 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-439-311-2
Query length: 333
Database: GenEmbl:*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 1486.240000


score_list:
```

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation ..
gb_pat:AR153100	-	50.00	88.70	1.1e+04	30	! AR153100 Sequence 102 from paten
gb_pat:AR153101	+	50.00	88.70	1.1e+04	30	! AR153101 Sequence 103 from paten
gb_pat:A93365	+	41.00	75.33	6.3e+04	50	! A93365 Sequence 4 from Patent WO
gb_pat:A12323	-	41.00	75.17	6.4e+04	51	! A12323 oligonucleotide. 12/1993
gb_pat:A12324	+	41.00	75.17	6.4e+04	51	! A12324 oligonucleotide. 12/1993
gb_pat:A12596	-	41.00	75.17	6.4e+04	51	! A12596 oligonucleotide. 1/1994
gb_pat:A12597	+	41.00	75.17	6.4e+04	51	! A12597 oligonucleotide. 1/1994
gb_pat:AX011428	+	40.00	72.99	8.5e+04	59	! AX011428 Sequence 105 from Paten
gb_pat:AR125926	+	40.00	72.86	8.6e+04	60	! AR125926 Sequence 268 from paten
gb_pat:I24293	+	40.00	72.86	8.6e+04	60	! I24293 Sequence 80 from patent U
gb_pat:AR013897	-	39.00	73.72	7.7e+04	47	! AR013897 Sequence 97 from patent
gb_pat:AR033851	-	39.00	73.72	7.7e+04	47	! AR033851 Sequence 97 from patent
gb_pat:AR042511	-	39.00	73.72	7.7e+04	47	! AR042511 Sequence 97 from patent
gb_pat:AR058391	-	39.00	73.72	7.7e+04	47	! AR058391 Sequence 97 from patent
gb_pat:AR088217	-	39.00	73.72	7.7e+04	47	! AR088217 Sequence 97 from patent
gb_pl:AF163970	-	39.00	72.93	8.6e+04	52	! AF163970 Saccharomyces cerevisia
gb_pat:AX156673	+	38.00	72.04	9.6e+04	51	! AX156673 Sequence 1 from Patent
gb_pat:AX158757	-	38.00	72.04	9.6e+04	51	! AX158757 Sequence 2085 from Pate
gb_pat:AX158758	-	38.00	72.04	9.6e+04	51	! AX158758 Sequence 2086 from Pate
gb_pat:AX158759	-	38.00	72.04	9.6e+04	51	! AX158759 Sequence 2087 from Pate
gb_pat:AR009607	-	38.00	71.74	9.9e+04	53	! AR009607 Sequence 3 from patent
gb_pat:AR036067	-	38.00	71.74	9.9e+04	53	! AR036067 Sequence 3 from patent
gb_pat:I27676	-	38.00	71.74	9.9e+04	53	! I27676 Sequence 3 from patent US
gb_pat:AX015223	+	38.00	70.77	1.1e+05	60	! AX015223 Sequence 1 from Patent
gb_pat:AR001228	-	37.00	76.88	5.2e+04	24	! AR001228 Sequence 6 from patent
gb_pat:AR008251	-	37.00	76.88	5.2e+04	24	! AR008251 Sequence 6 from patent
gb_pat:AR010178	-	37.00	76.88	5.2e+04	24	! AR010178 Sequence 6 from patent
gb_pat:AR064102	-	37.00	76.88	5.2e+04	24	! AR064102 Sequence 6 from patent
gb_pat:I38291	-	37.00	76.88	5.2e+04	24	! I38291 Sequence 6 from patent US
gb_pat:AR109123	-	37.00	72.70	8.8e+04	41	! AR109123 Sequence 82 from patent
gb_pat:E31328	+	37.00	71.31	1.0e+05	49	! E31328 Process for producing nov
gb_pat:AX158668	+	37.00	71.00	1.1e+05	51	! AX158668 Sequence 1996 from Pate
gb_pat:AX158963	+	37.00	71.00	1.1e+05	51	! AX158963 Sequence 2291 from Pate
gb_pat:AX158964	+	37.00	71.00	1.1e+05	51	! AX158964 Sequence 2292 from Pate
gb_pat:AX161669	+	37.00	71.00	1.1e+05	51	! AX161669 Sequence 4997 from Pate
gb_pat:A14709	+	37.00	70.70	1.1e+05	53	! A14709 Nucleotide sequence 15 fr
gb_pat:AR028022	-	37.00	69.73	1.3e+05	60	! AR028022 Sequence 4 from patent
gb_pat:AR101743	-	37.00	69.73	1.3e+05	60	! AR101743 Sequence 5 from patent
gb_pr:HUMTCRVJ65	+	37.00	69.73	1.3e+05	60	! L39603 Homo sapiens (PB.3.FA635)
gb_pat:A32410	-	36.00	73.35	8.1e+04	33	! A32410 Synthetic PAI-2 pDBP link
gb_pat:AX137684	+	36.00	73.12	8.4e+04	34	! AX137684 Sequence 17 from Patent
gb_pat:AX137685	-	36.00	73.12	8.4e+04	34	! AX137685 Sequence 18 from Pa
gb_pat:I08714	-	36.00	71.47	1.0e+05	42	! I08714 Sequence 2 from Paten
gb_pat:AX115993	-	36.00	69.95	1.2e+05	51	! AX115993 Sequence 1116 from
gb_ro:AF005617	+	36.00	69.95	1.2e+05	51	! AF005617 Mus musculus T cell

```
seq_name: gb_pat:AR153100

seq_documentation_block:
LOCUS       AR153100        30 bp    DNA              PAT       08-AUG-2001
DEFINITION  Sequence 102 from patent US 6235480.
ACCESSION   AR153100
VERSION     AR153100.1  GI:15120632
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Shultz,J.William, Lewis,M.K., Leippe,D., Mandrekar,M., Kephart,D.,
            Rhodes,R.Byron, Andrews,C.Ann, Hartnett,J.Robert, Gu,T.,
            Olson,R.J., Wood,K.V. and Welch,R.
  TITLE     Detection of nucleic acid hybrids
  JOURNAL   Patent: US 6235480-A 102 22-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        5 a      5 c      4 g     16 t
ORIGIN


alignment_scores:
            Quality:   50.00             Length:     10
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-439-311-2 x AR153100/rev   ..

 Align seg 1/1  to reverse of: AR153100  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
      30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1

seq_name: gb_pat:AR153101

seq_documentation_block:
LOCUS       AR153101        30 bp    DNA              PAT       08-AUG-2001
DEFINITION  Sequence 103 from patent US 6235480.
ACCESSION   AR153101
VERSION     AR153101.1  GI:15120633
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Shultz,J.William, Lewis,M.K., Leippe,D., Mandrekar,M., Kephart,D.,
            Rhodes,R.Byron, Andrews,C.Ann, Hartnett,J.Robert, Gu,T.,
            Olson,R.J., Wood,K.V. and Welch,R.
  TITLE     Detection of nucleic acid hybrids
  JOURNAL   Patent: US 6235480-A 103 22-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT       16 a      4 c      5 g      5 t
ORIGIN


alignment_scores:
            Quality:   50.00             Length:     10
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
```

```
 US-09-439-311-2 x AR153101   ..

 Align seg 1/1  to: AR153101  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
       1 CAAGATGGACAAAGTTTAAAAACAAGAACT 30

seq_name: gb_pat:A93365

seq_documentation_block:
LOCUS       A93365          50 bp    DNA             PAT       22-JAN-2000
DEFINITION  Sequence 4 from Patent WO9744451.
ACCESSION   A93365
VERSION     A93365.1  GI:6741628
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Paesen,G.C. and Nuttall,P.A.
  TITLE     VASOACTIVE AMINE BINDING MOLECULES
  JOURNAL   Patent: WO 9744451-A 4 27-NOV-1997;
            OXFORD VACS LTD (GB); PAESEN GUIDO CHRISTIAN (GB)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       11 a      8 c     14 g     17 t
ORIGIN


alignment_scores:
            Quality:   41.00             Length:     10
              Ratio:  4.100               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 70.000

alignment_block:
 US-09-439-311-2 x A93365   ..

 Align seg 1/1  to: A93365  from: 1  to: 50

     262 SerAspGlyAspGluAsnGlySerLeuIle 271
         ||||||||||||:::::|||||||||:::
       7 AGTGATGGTGATGATGATGGATCCCTTCTG 36

seq_name: gb_pat:A12323

seq_documentation_block:
LOCUS       A12323          51 bp    DNA             PAT       06-DEC-1993
DEFINITION  oligonucleotide.
ACCESSION   A12323
VERSION     A12323.1  GI:491330
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   .
  TITLE     HYBRID PROTEINS OR POLYPEPTIDES
  JOURNAL   Patent: WO 8802757-A 24 21-APR-1988;
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       15 a     27 c      7 g      2 t
ORIGIN


alignment_scores:
            Quality:   41.00             Length:     17
              Ratio:  3.154               Gaps:      0
 Percent Similarity:  76.471   Percent Identity: 47.059

alignment_block:
 US-09-439-311-2 x A12323/rev   ..

 Align seg 1/1  to reverse of: A12323  from: 1  to: 51

     300 GlyArgGlyIleLysIleThrGlySerIleGlyValGlyAlaGlyIleLe 316
         |||   |||:::   :::      ::::::|||||||||:::|||||||
      51 GGGATCGGGGTTGGCGTTGGGGTTGGCGTTGGGGTTGGCGTTGGGATCCT 2

     316 u 316
         |
       1 C 1

seq_name: gb_pat:A12324

seq_documentation_block:
LOCUS       A12324          51 bp    DNA             PAT       06-DEC-1993
DEFINITION  oligonucleotide.
ACCESSION   A12324
VERSION     A12324.1  GI:489519
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   .
  TITLE     HYBRID PROTEINS OR POLYPEPTIDES
  JOURNAL   Patent: WO 8802757-A 25 21-APR-1988;
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        2 a      7 c     27 g     15 t
ORIGIN


alignment_scores:
            Quality:   41.00             Length:     17
              Ratio:  3.154               Gaps:      0
 Percent Similarity:  76.471   Percent Identity: 47.059

alignment_block:
 US-09-439-311-2 x A12324   ..

 Align seg 1/1  to: A12324  from: 1  to: 51

   . 300 GlyArgGlyIleLysIleThrGlySerIleGlyValGlyAlaGlyIleLe 316
         |||   |||:::   :::      ::::::|||||||||:::|||||||
       1 GGGATCGGGGTTGGCGTTGGGGTTGGCGTTGGGGTTGGCGTTGGGATCCT 50

     316 u 316
         |
      51 C 51

seq_name: gb_pat:A12596

seq_documentation_block:
LOCUS       A12596          51 bp    DNA             PAT       05-JAN-1994
DEFINITION  oligonucleotide.
ACCESSION   A12596
VERSION     A12596.1  GI:491421
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   .
  TITLE     RECOMBINANT VIRUS
  JOURNAL   Patent: WO 8701386-A 12 12-MAR-1987;
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
```

```
                     /db_xref="taxon:32630"
BASE COUNT       15 a     27 c      7 g      2 t
ORIGIN


alignment_scores:
            Quality:   41.00             Length:     17
              Ratio:  3.154               Gaps:      0
 Percent Similarity:  76.471   Percent Identity: 47.059

alignment_block:
 US-09-439-311-2 x A12596/rev   ..

 Align seg 1/1  to reverse of: A12596  from: 1  to: 51

    300 GlyArgGlyIleLysIleThrGlySerIleGlyValGlyAlaGlyIleLe 316
        |||   |||:::   :::      ::::::||||||||||:::|||||||
     51 GGGATCGGGGTTGGCGTTGGGGTTGGCGTTGGGGTTGGCGTTGGGATCCT 2

    316 u 316
        |
      1 C 1

seq_name: gb_pat:A12597

seq_documentation_block:
LOCUS       A12597         51 bp    DNA             PAT       05-JAN-1994
DEFINITION  oligonucleotide.
ACCESSION   A12597
VERSION     A12597.1  GI:489543
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   .
  TITLE     RECOMBINANT VIRUS
  JOURNAL   Patent: WO 8701386-A 13 12-MAR-1987;
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        2 a      7 c     27 g     15 t
ORIGIN


alignment_scores:
            Quality:   41.00             Length:     17
              Ratio:  3.154               Gaps:      0
 Percent Similarity:  76.471   Percent Identity: 47.059

alignment_block:
 US-09-439-311-2 x A12597   ..

 Align seg 1/1  to: A12597  from: 1  to: 51

    300 GlyArgGlyIleLysIleThrGlySerIleGlyValGlyAlaGlyIleLe 316
        |||   |||:::   :::      ::::::||||||||||:::|||||||
      1 GGGATCGGGGTTGGCGTTGGGGTTGGCGTTGGGGTTGGCGTTGGGATCCT 50

    316 u 316
        |
     51 C 51

seq_name: gb_pat:AX011428

seq_documentation_block:
LOCUS       AX011428       59 bp    DNA             PAT       06-SEP-2000
DEFINITION  Sequence 105 from Patent WO9955907.
ACCESSION   AX011428
VERSION     AX011428.1  GI:9997978
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Koetter,P., Entian,K.D. and Diu-Hercend,A.
  TITLE     Method for screening antimycotic substances using essential genes
            from s. Cerevisiae
  JOURNAL   Patent: WO 9955907-A 105 04-NOV-1999;
            KOETTER PETER (DE); ENTIAN KARL DIETER (DE); DIU HERCEND ANITA
            (FR); HOECHST MARION ROUSSEL INC (FR)
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer YMR290c-S1"
BASE COUNT        8 a     12 c     15 g     24 t
ORIGIN


alignment_scores:
            Quality:   40.00             Length:     15
              Ratio:  3.077               Gaps:      0
 Percent Similarity:  86.667   Percent Identity: 53.333

alignment_block:
 US-09-439-311-2 x AX011428   ..

 Align seg 1/1  to: AX011428  from: 1  to: 59

    163 SerLysIleGlyValThrArgPheGluThrGlySerGlnSerPhe 177
        :::::::::|||:::   ||||||:::|||||||||   ||||||
      9 ACGTCTTTTGGTATTTGGCGTTTTTCCACTGGCAGCTGAAGCTTC 53

seq_name: gb_pat:AR125926

seq_documentation_block:
LOCUS       AR125926       60 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 268 from patent US 6177557.
ACCESSION   AR125926
VERSION     AR125926.1  GI:14111988
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Janjic,N., Gold,L. and Tasset,D.
  TITLE     High affinity ligands of basic fibroblast growth factor and
            thrombin
  JOURNAL   Patent: US 6177557-A 268 23-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       10 a     11 c     29 g     10 t
ORIGIN


alignment_scores:
            Quality:   40.00             Length:     19
              Ratio:  2.667               Gaps:      0
 Percent Similarity:  78.947   Percent Identity: 42.105

alignment_block:
 US-09-439-311-2 x AR125926   ..

 Align seg 1/1  to: AR125926  from: 1  to: 60

    204 ThrSerValGlyThrGlyLeuGlyAlaLeuAlaGluGluIleAsnArgAs 220
        |||:::   |||:::||||||   |||||||||:::   ::::::   ::
      4 ACCGCGGAGGGCGTAGGGTTGGAGGCGTTGGCCGATGTGGTAGGCACGGA 53

    220 nAlaAsp 222
        :::::|||
     54 CTCGGAT 60
```

```
seq_name: gb_pat:I24293

seq_documentation_block:
LOCUS       I24293          60 bp    DNA             PAT       07-OCT-1996
DEFINITION  Sequence 80 from patent US 5543293.
ACCESSION   I24293
VERSION     I24293.1  GI:1604163
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Gold,L. and Tasset,D.
  TITLE     DNA ligands of thrombin
  JOURNAL   Patent: US 5543293-A 80 06-AUG-1996;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       10 a     11 c     29 g     10 t
ORIGIN


alignment_scores:
            Quality:   40.00             Length:     19
              Ratio:  2.667               Gaps:      0
 Percent Similarity:  78.947   Percent Identity: 42.105

alignment_block:
 US-09-439-311-2 x I24293   ..

 Align seg 1/1  to: I24293  from: 1  to: 60

    204 ThrSerValGlyThrGlyLeuGlyAlaLeuAlaGluGluIleAsnArgAs 220
        |||:::   |||:::||||||   ||||||||||:::   ::::::   ::
      4 ACCGCGGAGGGCGTAGGGTTGGAGGCGTTGGCCGATGTGGTAGGCACGGA 53

    220 nAlaAsp 222
        ::::|||
     54 CTCGGAT 60

seq_name: gb_pat:AR013897

seq_documentation_block:
LOCUS       AR013897        47 bp    DNA             PAT       05-DEC-1998
DEFINITION  Sequence 97 from patent US 5773218.
ACCESSION   AR013897
VERSION     AR013897.1  GI:3971351
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Gallatin,W.Michael and Vazeux,R.
  TITLE     Method to identify compounds which modulate ICAM-related protein
            interactions
  JOURNAL   Patent: US 5773218-A 97 30-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT        9 a     21 c      7 g     10 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     12
              Ratio:  3.900               Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AR013897/rev   ..

 Align seg 1/1  to reverse of: AR013897  from: 1  to: 47

    169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
        |||:::||||||||||||||      |||:::||||||
     44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: gb_pat:AR033851

seq_documentation_block:
LOCUS       AR033851        47 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 97 from patent US 5869262.
ACCESSION   AR033851
VERSION     AR033851.1  GI:5949456
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Gallatin,W.Michael and Vazeux,R.
  TITLE     Method for monitoring an inflammatory disease state by detecting
            circulating ICAM-R
  JOURNAL   Patent: US 5869262-A 97 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT        9 a     21 c      7 g     10 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     12
              Ratio:  3.900               Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AR033851/rev   ..

 Align seg 1/1  to reverse of: AR033851  from: 1  to: 47

    169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
        |||:::||||||||||||||      |||:::||||||
     44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: gb_pat:AR042511

seq_documentation_block:
LOCUS       AR042511        47 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 97 from patent US 5811517.
ACCESSION   AR042511
VERSION     AR042511.1  GI:5963007
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Gallatin,W.Michael and Vazeux,R.
  TITLE     ICAM-related protein variants
  JOURNAL   Patent: US 5811517-A 97 22-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT        9 a     21 c      7 g     10 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     12
              Ratio:  3.900               Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AR042511/rev   ..

 Align seg 1/1  to reverse of: AR042511  from: 1  to: 47
```

```
     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9
```

```
seq_name: gb_pat:AR058391

seq_documentation_block:
LOCUS       AR058391       47 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 97 from patent US 5837822.
ACCESSION   AR058391
VERSION     AR058391.1  GI:5983968
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Gallatin,W.Michael and Vazeux,R.
  TITLE     Humanized antibodies specific for ICAM related protein
  JOURNAL   Patent: US 5837822-A 97 17-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT        9 a     21 c      7 g     10 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     12
              Ratio:  3.900               Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AR058391/rev   ..

 Align seg 1/1  to reverse of: AR058391  from: 1  to: 47
```

```
     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9
```

```
seq_name: gb_pat:AR088217

seq_documentation_block:
LOCUS       AR088217       47 bp    DNA             PAT       07-SEP-2000
DEFINITION  Sequence 97 from patent US 5989843.
ACCESSION   AR088217
VERSION     AR088217.1  GI:10014980
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Gallatin,W.Michael and Vazeux,R.
  TITLE     Methods for identifying modulators of protein kinase C
            phosphorylation of ICAM-related protein
  JOURNAL   Patent: US 5989843-A 97 23-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT        9 a     21 c      7 g     10 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     12
              Ratio:  3.900               Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AR088217/rev   ..

 Align seg 1/1  to reverse of: AR088217  from: 1  to: 47
```

```
     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 16, 2002, 23:26:24 ; Search time 1531.8 Seconds
(without alignments)
10759.030 Million cell updates/sec

Title: US-09-439-311-1
Perfect score: 999
Sequence: 1 attaacacaaatgttgcagc.....ttaaaaatgatggtagagat 999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 586436

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.		Score	Query Match	Length	DB	ID	Description
c	1	26.8	2.7	30	6	AR153100	AR153100 Sequence
	2	26.8	2.7	30	6	AR153101	AR153101 Sequence
	3	21.4	2.1	60	3	AF320167	AF320167 Drosophil
c	4	20.8	2.1	24	6	AR001228	AR001228 Sequence
c	5	20.8	2.1	24	6	AR008251	AR008251 Sequence
c	6	20.8	2.1	24	6	AR010178	AR010178 Sequence
c	7	20.8	2.1	24	6	AR064102	AR064102 Sequence
c	8	20.8	2.1	24	6	I38291	I38291 Sequence 6
	9	20.8	2.1	44	6	AX052949	AX052949 Sequence
c	10	20.8	2.1	58	6	AR061278	AR061278 Sequence
	11	20.6	2.1	60	6	A80405	A80405 Sequence 17
	12	20.4	2.0	51	6	AX160493	AX160493 Sequence
	13	20.4	2.0	51	6	AX160494	AX160494 Sequence
	14	20.4	2.0	57	3	AF320169	AF320169 Drosophil
	15	20.4	2.0	57	3	AF320170	AF320170 Drosophil
	16	20.4	2.0	57	3	AF320171	AF320171 Drosophil
	17	20.4	2.0	57	3	AF320172	AF320172 Drosophil
	18	20.4	2.0	57	3	AF320173	AF320173 Drosophil
	19	20.4	2.0	57	3	AF320174	AF320174 Drosophil
	20	20.4	2.0	57	3	AF320175	AF320175 Drosophil
	21	20.4	2.0	57	3	AF320176	AF320176 Drosophil
	22	20.4	2.0	57	3	AF320177	AF320177 Drosophil
	23	20.4	2.0	57	3	AF320178	AF320178 Drosophil
	24	20.4	2.0	57	3	AF320179	AF320179 Drosophil
	25	20.2	2.0	52	6	A69988	A69988 Sequence 19
c	26	20	2.0	50	6	A17107	A17107 Oligonucleo
c	27	20	2.0	50	6	AR027492	AR027492 Sequence
	28	20	2.0	57	6	AR077133	AR077133 Sequence
	29	20	2.0	57	6	AR102800	AR102800 Sequence
	30	20	2.0	57	6	I21136	I21136 Sequence 4
	31	20	2.0	60	6	A38623	A38623 Sequence 16
	32	20	2.0	60	6	AR040717	AR040717 Sequence
c	33	19.8	2.0	51	6	AX165291	AX165291 Sequence
c	34	19.8	2.0	59	6	AX011435	AX011435 Sequence
	35	19.6	2.0	29	6	A20557	A20557 oligonucleo
c	36	19.6	2.0	43	6	I11627	I11627 Sequence 12
c	37	19.6	2.0	51	6	AX160235	AX160235 Sequence
	38	19.6	2.0	55	14	HIV1U45041	U45041 Human immun
	39	19.6	2.0	57	6	A32988	A32988 Synthetic P
c	40	19.6	2.0	58	6	AR011305	AR011305 Sequence
c	41	19.6	2.0	58	6	I17943	I17943 Sequence 17
c	42	19.4	1.9	37	6	AR063204	AR063204 Sequence
	43	19.4	1.9	49	14	ALAM03	V00054 Alfalfa mos
	44	19.4	1.9	51	6	AX117493	AX117493 Sequence
c	45	19.4	1.9	51	6	AX162140	AX162140 Sequence

ALIGNMENTS

RESULT 1
AR153100/c
LOCUS AR153100 30 bp DNA PAT 08-AUG-2001
DEFINITION Sequence 102 from patent US 6235480.
ACCESSION AR153100
VERSION AR153100.1 GI:15120632
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Shultz,J.William, Lewis,M.K., Leippe,D., Mandrekar,M., Kephart,D., Rhodes,R.Byron, Andrews,C.Ann, Hartnett,J.Robert, Gu,T., Olson,R.J., Wood,K.V. and Welch,R.
TITLE Detection of nucleic acid hybrids
JOURNAL Patent: US 6235480-A 102 22-MAY-2001;
FEATURES Location/Qualifiers
source 1. .30
/organism="unknown"
BASE COUNT 5 a 5 c 4 g 16 t

```
ORIGIN


  Query Match             2.7%;  Score 26.8;  DB 6;  Length 30;
  Best Local Similarity   93.3%;  Pred. No. 1.4e+05;
  Matches   28;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    289 caagatggtcaaagcttaaaaacaagaact 318
          |||||||| ||||| ||||||||||||||||
Db     30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1


RESULT   2
AR153101
LOCUS       AR153101       30 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 103 from patent US 6235480.
ACCESSION   AR153101
VERSION     AR153101.1  GI:15120633
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Shultz,J.William, Lewis,M.K., Leippe,D., Mandrekar,M., Kephart,D.,
            Rhodes,R.Byron, Andrews,C.Ann, Hartnett,J.Robert, Gu,T.,
            Olson,R.J., Wood,K.V. and Welch,R.
  TITLE     Detection of nucleic acid hybrids
  JOURNAL   Patent: US 6235480-A 103 22-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT       16 a      4 c      5 g      5 t
ORIGIN


  Query Match             2.7%;  Score 26.8;  DB 6;  Length 30;
  Best Local Similarity   93.3%;  Pred. No. 1.4e+05;
  Matches   28;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    289 caagatggtcaaagcttaaaaacaagaact 318
          |||||||| ||||| ||||||||||||||||
Db      1 CAAGATGGACAAAGTTTAAAAACAAGAACT 30


RESULT   3
AF320167
LOCUS       AF320167       60 bp    DNA             INV       23-APR-2001
DEFINITION  Drosophila pseudoobscura strain Abajo36 bicoid (bcd) gene, partial
            cds.
ACCESSION   AF320167
VERSION     AF320167.1  GI:13752322
KEYWORDS    .
SOURCE      Drosophila pseudoobscura.
  ORGANISM  Drosophila pseudoobscura
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Noor,M.A., Kliman,R.M. and Machado,C.A.
  TITLE     Evolutionary history of microsatellites in the obscura group of
            Drosophila
  JOURNAL   Mol. Biol. Evol. 18 (4), 551-556 (2001)
  MEDLINE   21165327
   PUBMED   11264406
REFERENCE   2  (bases 1 to 60)
  AUTHORS   Noor,M.A.F., Kliman,R.M. and Machado,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2000) Genetics, Rutgers University, 604 Allison
            Rd., Piscataway, NJ 08854, USA
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Drosophila pseudoobscura"
                     /strain="Abajo36"
                     /db_xref="taxon:7237"
     gene            <1. .>60
                     /gene="bcd"
     exon            <1. .>60
                     /gene="bcd"
                     /number=3
     repeat_region   1. .60
                     /note="microsatellite"
                     /rpt_type=tandem
     CDS             <1. .>60
                     /gene="bcd"
                     /codon_start=1
                     /product="bicoid"
                     /protein_id="AAK38619.1"
                     /db_xref="GI:13752323"
                     /translation="XAQFFQTQQQQQLHQQQQQQ"
BASE COUNT       21 a     19 c     12 g      7 t      1 others
ORIGIN


  Query Match             2.1%;  Score 21.4;  DB 3;  Length 60;
  Best Local Similarity   59.3%;  Pred. No. 1.4e+06;
  Matches   35;  Conservative    1;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     84 cagttcaggtcttagaatcaactccgcagcagatgatgcttcagggatggcgatagcag 142
          :||  ||| ||||  |  ||   | |||||||  | | | ||||     || | |||||
Db      2 BAGCACAGTTCTTTCAGACACAACAGCAGCAGCAGCTCCATCAGCAGCAGCAACAGCAG 60


RESULT   4
AR001228/c
LOCUS       AR001228       24 bp    DNA             PAT       04-DEC-1998
DEFINITION  Sequence 6 from patent US 5738995.
ACCESSION   AR001228
VERSION     AR001228.1  GI:3963295
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Wu,L., Coombs,J., Malmstrom,S.L. and Glass,M.J.
  TITLE     Inosine-containing probes for detecting E.coli 0157:H7
  JOURNAL   Patent: US 5738995-A 6 14-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        6 a      5 c      5 g      8 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 6;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 1.9e+06;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     91 ggtcttagaatcaactccgcagca 114
          |||||||||||| ||||| ||||||
Db     24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   5
AR008251/c
LOCUS       AR008251       24 bp    DNA             PAT       04-DEC-1998
DEFINITION  Sequence 6 from patent US 5753444.
ACCESSION   AR008251
VERSION     AR008251.1  GI:3967360
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Wu,L., Coombs,J., Malmstrom,S.L. and Glass,M.J.
  TITLE     Methods and kits using inosine-containing probes for discriminating
            variant nucleic acid sequences
```

```
  JOURNAL   Patent: US 5753444-A 6 19-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        6 a      5 c      5 g      8 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 6;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 1.9e+06;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   6
AR010178/c
LOCUS       AR010178       24 bp    DNA             PAT       04-DEC-1998
DEFINITION  Sequence 6 from patent US 5756701.
ACCESSION   AR010178
VERSION     AR010178.1  GI:3968983
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Wu,L., Coombs,J., Malmstrom,S.L. and Glass,M.J.
  TITLE     Specific oligonucleotide primer pairs and probes for discriminating
            specific analytes
  JOURNAL   Patent: US 5756701-A 6 26-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        6 a      5 c      5 g      8 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 6;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 1.9e+06;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   7
AR064102/c
LOCUS       AR064102       24 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 6 from patent US 5846783.
ACCESSION   AR064102
VERSION     AR064102.1  GI:5993410
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Wu,L., Coombs,J., Malmstrom,S.L. and Glass,M.J.
  TITLE     Methods and apparatus for preparing, amplifying, and discriminating
            multiple analytes
  JOURNAL   Patent: US 5846783-A 6 08-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        6 a      5 c      5 g      8 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 6;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 1.9e+06;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   8
I38291/c
LOCUS       I38291         24 bp    DNA             PAT       13-MAY-1997
DEFINITION  Sequence 6 from patent US 5612473.
ACCESSION   I38291
VERSION     I38291.1  GI:2086281
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Wu,L., Coombs,J., Malmstrom,S.L. and Glass,M.J.
  TITLE     Methods, kits and solutions for preparing sample material for
            nucleic acid amplification
  JOURNAL   Patent: US 5612473-A 6 18-MAR-1997;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        6 a      5 c      5 g      8 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 6;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 1.9e+06;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   9
AX052949
LOCUS       AX052949       44 bp    DNA             PAT       12-JAN-2001
DEFINITION  Sequence 55 from Patent WO0071755.
ACCESSION   AX052949
VERSION     AX052949.1  GI:12227051
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Kwagh,J.G., Macklin,J.J., Mitsis,P.G. and Ulmer,K.M.
  TITLE     Method for sequency and characterizing polymeric biomolecules using
            aptamers and a method for producing aptamers
  JOURNAL   Patent: WO 0071755-A 55 30-NOV-2000;
            Praelux Incorporated (US)
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic DNA sequence"
BASE COUNT        9 a      6 c     19 g     10 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 6;  Length 44;
  Best Local Similarity   70.0%;  Pred. No. 1.8e+06;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy     747 tgctatcaatggggttgttataggtaaggttgattattca 786
           ||  | || |||||| | ||| |||| |||||   | |||
Db       4 TGACACCACTGGGGTGGGTATGGGTAGGGTTGTGGAATCA 43


RESULT  10
AR061278/c
```

```
LOCUS       AR061278       58 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 7 from patent US 5843650.
ACCESSION   AR061278
VERSION     AR061278.1  GI:5988969
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Segev,D.
  TITLE     Nucleic acid detection and amplification by chemical linkage of
            oligonucleotides
  JOURNAL   Patent: US 5843650-A 7 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="unknown"
BASE COUNT       15 a     18 c     13 g     12 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 6;  Length 58;
  Best Local Similarity   60.7%;  Pred. No. 1.8e+06;
  Matches   34;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy    789 tggtgatgagaatggttctttaatttcagctatcaatgctgtaaaagatacaactg 844
          || || |||| ||||  ||    ||   |||  | |||| | ||  | ||||  ||
Db     58 TGATGGTGAGGATGGGCCTCCGGTTCATGCTGCCCATGCAGGAACTGTTACACATG 3


RESULT  11
A80405
LOCUS       A80405         60 bp    DNA             PAT       21-JAN-2000
DEFINITION  Sequence 17 from Patent WO9951771.
ACCESSION   A80405
VERSION     A80405.1  GI:6731293
KEYWORDS    .
SOURCE      Archaeoglobus fulgidus.
  ORGANISM  Archaeoglobus fulgidus
            Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
            Archaeoglobus.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Jansen,R. and Schouls,L.M.
  TITLE     A METHOD OF INTERSTRAIN DIFFERENTIATION OF BACTERIA
  JOURNAL   Patent: WO 9951771-A 17 14-OCT-1999;
            EMBDEN JOHANNES DIRK ANTHONIE (NL); JANSEN RUDOLPH (NL)
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Archaeoglobus fulgidus"
                     /db_xref="taxon:2234"
BASE COUNT       20 a     11 c      9 g     20 t
ORIGIN


  Query Match             2.1%;  Score 20.6;  DB 6;  Length 60;
  Best Local Similarity   62.7%;  Pred. No. 2e+06;
  Matches   32;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    561 aaactacaacggtatcgaagattttaaatttgatagtgtagtgatttctac 611
          | || | || || || |||   ||| |||   ||  ||   ||||||| ||
Db     10 AGACCAAAATGGGATTGAAATCTTTCAATCCCATTTTGGTCTGATTTCAAC 60


RESULT  12
AX160493
LOCUS       AX160493       51 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 3821 from Patent WO0140521.
ACCESSION   AX160493
VERSION     AX160493.1  GI:14541824
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3821 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (3822 is other entry)
                     Accession number cg43921050"
BASE COUNT       19 a      5 c      8 g     19 t
ORIGIN


  Query Match             2.0%;  Score 20.4;  DB 6;  Length 51;
  Best Local Similarity   65.2%;  Pred. No. 2.2e+06;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    858 taaagatgaaaatggtaaacttgttcttacttcggccgatggtaga 903
          ||||||||||  | ||| |   || ||  ||  |  ||  |||||
Db      4 TAAAGATGAAACAGTTAAGCCAATTTTTTTTTTTGAAGAATGTAGA 49


RESULT  13
AX160494
LOCUS       AX160494       51 bp    DNA             PAT       22-JUN-2001
DEFINITION  Sequence 3822 from Patent WO0140521.
ACCESSION   AX160494
VERSION     AX160494.1  GI:14541825
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3822 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (3821 is other entry)
                     Accession number cg43921050"
BASE COUNT       18 a      5 c      8 g     20 t
ORIGIN


  Query Match             2.0%;  Score 20.4;  DB 6;  Length 51;
  Best Local Similarity   65.2%;  Pred. No. 2.2e+06;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    858 taaagatgaaaatggtaaacttgttcttacttcggccgatggtaga 903
          ||||||||||  | ||| |   || ||  ||  |  ||  |||||
Db      4 TAAAGATGAAACAGTTAAGCCATTTTTTTTTTTGAAGAATGTAGA 49


RESULT  14
AF320169
LOCUS       AF320169       57 bp    DNA             INV       23-APR-2001
DEFINITION  Drosophila pseudoobscura strain Mather10 bicoid (bcd) gene, partial
            cds.
ACCESSION   AF320169
VERSION     AF320169.1  GI:13752326
KEYWORDS    .
SOURCE      Drosophila pseudoobscura.
  ORGANISM  Drosophila pseudoobscura
```

```
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Noor,M.A., Kliman,R.M. and Machado,C.A.
  TITLE     Evolutionary history of microsatellites in the obscura group of
            Drosophila
  JOURNAL   Mol. Biol. Evol. 18 (4), 551-556 (2001)
  MEDLINE   21165327
   PUBMED   11264406
REFERENCE   2  (bases 1 to 57)
  AUTHORS   Noor,M.A.F., Kliman,R.M. and Machado,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2000) Genetics, Rutgers University, 604 Allison
            Rd., Piscataway, NJ 08854, USA
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Drosophila pseudoobscura"
                     /strain="Mather10"
                     /db_xref="taxon:7237"
     gene            <1. .>57
                     /gene="bcd"
     exon            <1. .>57
                     /gene="bcd"
                     /number=3
     repeat_region   1. .57
                     /note="microsatellite"
                     /rpt_type=tandem
     CDS             <1. .>57
                     /gene="bcd"
                     /codon_start=1
                     /product="bicoid"
                     /protein_id="AAK38621.1"
                     /db_xref="GI:13752327"
                     /translation="AQFFQTQQQQQLHQQQQQQ"
BASE COUNT       19 a     19 c     12 g      7 t
ORIGIN


  Query Match             2.0%;  Score 20.4;  DB 3;  Length 57;
  Best Local Similarity   61.1%;  Pred. No. 2.2e+06;
  Matches   33;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      89 caggtcttagaatcaactccgcagcagatgatgcttcagggatggcgatagcag 142
           ||| ||||  |  ||   | |||||||  | | | ||||     || | |||||
Db       4 CAGTTCTTTCAGACACAACAGCAGCAGCAGCTCCATCAGCAGCAGCAACAGCAG 57


RESULT  15
AF320170
LOCUS       AF320170       57 bp    DNA             INV       23-APR-2001
DEFINITION  Drosophila pseudoobscura strain Mather17 bicoid (bcd) gene, partial
            cds.
ACCESSION   AF320170
VERSION     AF320170.1  GI:13752328
KEYWORDS    .
SOURCE      Drosophila pseudoobscura.
  ORGANISM  Drosophila pseudoobscura
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Noor,M.A., Kliman,R.M. and Machado,C.A.
  TITLE     Evolutionary history of microsatellites in the obscura group of
            Drosophila
  JOURNAL   Mol. Biol. Evol. 18 (4), 551-556 (2001)
  MEDLINE   21165327
   PUBMED   11264406
REFERENCE   2  (bases 1 to 57)
  AUTHORS   Noor,M.A.F., Kliman,R.M. and Machado,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2000) Genetics, Rutgers University, 604 Allison
            Rd., Piscataway, NJ 08854, USA
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Drosophila pseudoobscura"
                     /strain="Mather17"
                     /db_xref="taxon:7237"
     gene            <1. .>57
                     /gene="bcd"
     exon            <1. .>57
                     /gene="bcd"
                     /number=3
     repeat_region   1. .57
                     /note="microsatellite"
                     /rpt_type=tandem
     CDS             <1. .>57
                     /gene="bcd"
                     /codon_start=1
                     /product="bicoid"
                     /protein_id="AAK38622.1"
                     /db_xref="GI:13752329"
                     /translation="AQFFQTQQQQQLHQQQQQQ"
BASE COUNT       19 a     19 c     12 g      7 t
ORIGIN


  Query Match             2.0%;  Score 20.4;  DB 3;  Length 57;
  Best Local Similarity   61.1%;  Pred. No. 2.2e+06;
  Matches   33;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      89 caggtcttagaatcaactccgcagcagatgatgcttcagggatggcgatagcag 142
           ||| ||||  |  ||   | |||||||  | | | ||||     || | |||||
Db       4 CAGTTCTTTCAGACACAACAGCAGCAGCAGCTCCATCAGCAGCAGCAACAGCAG 57


Search completed: April 17, 2002, 01:49:31
Job time: 8587 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 17, 2002, 01:18:44 ; Search time 172.86 Seconds
(without alignments)
4954.691 Million cell updates/sec

Title: US-09-439-311-1
Perfect score: 999
Sequence: 1 attaacacaaatgttgcagc.....ttaaaaatgatggtagagat 999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1026190

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	26.8	2.7	30	21	AAA86891	Probe to Campyloba
	2	26.8	2.7	30	21	AAA86892	Probe to Campyloba
c	3	26.8	2.7	30	21	AAA93188	Campylobacter jeju
	4	26.8	2.7	30	21	AAA93190	Campylobacter jeju
c	5	21.8	2.2	50	20	AAX52169	Synthetic plasmid
	6	21.4	2.1	51	19	AAV04223	Human cardiac trop
c	7	21.4	2.1	58	19	AAV04224	Human cardiac trop
	8	21.2	2.1	54	21	AAA97417	Pea wild-type pra2
	9	21	2.1	27	21	AAA27148	Campylobacter coli
c	10	21	2.1	33	21	AAA27149	Campylobacter coli
c	11	20.8	2.1	24	18	AAT60508	Primer CFO4R.4. S
c	12	20.8	2.1	24	19	AAV31446	Campylobacter nucl
c	13	20.8	2.1	24	19	AAV25942	Oligonucleotide PC
c	14	20.8	2.1	24	19	AAV20847	Campylobacter CFO4
	15	20.8	2.1	44	22	AAF16711	dGMP-specific apta
	16	20.8	2.1	50	18	AAV76100	Staphylococcus aur
c	17	20.8	2.1	51	22	AAH89291	Human coding sequ
c	18	20.8	2.1	58	19	AAV15200	Human serrate 2 PC
	19	20.6	2.1	47	21	AAZ67004	Human map-related
c	20	20.6	2.1	51	18	AAT70191	Primer SEQ ID NO:2
c	21	20.6	2.1	51	20	AAX16831	Human delta-1 gene
c	22	20.4	2.0	47	21	AAZ66354	Human map-related
	23	20.4	2.0	47	21	AAZ68687	Human map-related
	24	20.4	2.0	50	16	AAT25074	Human gene signatu
	25	20.2	2.0	50	22	AAH25534	PCR primer used to
c	26	20.2	2.0	50	22	AAH25535	PCR primer used to
	27	20.2	2.0	59	21	AAC24094	Human secreted pro
	28	20.2	2.0	60	16	AAT21262	Human gene signatu
c	29	20	2.0	39	19	AAV05387	Primer pAS3BACREV
c	30	20	2.0	39	21	AAA55959	Sequencing and PCR
	31	20	2.0	44	21	AAZ55398	Neisseria species
c	32	20	2.0	52	20	AAX16314	Human delta-2 prim
	33	20	2.0	57	20	AAZ07541	ST receptor peptid
	34	20	2.0	60	15	AAQ66366	McPC603 V-min gene
	35	20	2.0	60	17	AAT38741	Moraxella outer me
c	36	19.8	2.0	45	22	AAF30242	Dystrophin exon 3
	37	19.8	2.0	47	21	AAZ69525	Human map-related
c	38	19.8	2.0	59	21	AAZ96931	S. cerevisiae gene
c	39	19.6	2.0	36	14	AAQ40036	Sequence of revers
c	40	19.6	2.0	43	16	AAQ89640	Reverse primer 99R
c	41	19.6	2.0	47	21	AAZ66795	Human map-related
c	42	19.6	2.0	47	21	AAZ67933	Human map-related
	43	19.6	2.0	57	13	AAQ23824	Primer HuClambdaFO
c	44	19.6	2.0	58	13	AAQ35636	SIV env primer SIV
c	45	19.6	2.0	58	14	AAQ35359	PCR primer SIVENV4

ALIGNMENTS

RESULT 1
AAA86891/c
ID AAA86891 standard; DNA; 30 BP.
XX
AC AAA86891;
XX
DT 15-JAN-2001 (first entry)
XX
DE Probe to Campylobacter jejuni.
XX
KW Detection; nucleic acid hybrid; depolymerisation; analysis; SNP;
KW single nucleotide polymorphism; identification; viral load; probe;
KW genotyping; medical marker diagnostic; primer; target; mutation;
KW genetic disease; ss.
XX
OS Campylobacter jejuni.
XX
PN WO200049180-A1.
XX
PD 24-AUG-2000.
XX
PF 18-FEB-2000; 2000WO-US04242.
XX
PR 18-FEB-1999; 99US-0252436.
PR 21-JUL-1999; 99US-0358972.
PR 25-AUG-1999; 99US-0383316.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Shultz JW, Lewis MK, Leippe D, Mandrekar M, Kephart D, Rhodes RB;
PI Andrews CA, Hartnett JR, Gu T, Olson RJ, Wood KV, Welch R;
XX
DR WPI; 2000-565377/52.
XX

PT Determining presence or absence of a predetermined endogenous nucleic
PT acid sequence by using an enzyme that depolymerizes the 3' end of an
PT oligonucleotide probe hybridized to a target sequence to release
PT identifier nucleotides -
XX
PS Example; Page 321; 389pp; English.
XX
CC The present invention describes a method (M1) for determining the
CC presence or absence of a predetermined endogenous nucleic acid target
CC sequence (ENAT). The method comprises hybridising a probe having an
CC identifier nucleotide (IN) with ENAT which is treated with an enzyme
CC that depolymerises the 3' end of hybridised NA to release the INs.
CC M1 is used for determining the number of known sequence repeats present
CC in a nucleic acid target sequence in a nucleic acid sample. The method
CC is also useful for determining whether a nucleic acid target sequence in
CC a sample is an allele from a homozygous or heterozygous locus. The
CC method is also useful for detection of mutations, translocations and
CC SNPs in nucleic acids (including those associated with genetic disease),
CC determination of viral load, species identification, sample
CC contamination, and analysis of forensic samples. AAA86791 to AAA87079
CC and AAB12817 represent sequence which are used in the exemplification of
CC the present invention.
CC N.B. There is a discrepancy between the SEQ ID NO: and sequences given
CC in the examples, and the SEQ ID NO: and sequences given in the sequence
CC listing from the present invention.
XX
SQ Sequence 30 BP; 5 A; 5 C; 4 G; 16 T; 0 other;

Query Match 2.7%; Score 26.8; DB 21; Length 30;
Best Local Similarity 93.3%; Pred. No. 1.3e+03;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 289 caagatggtcaaagcttaaaaacaagaact 318
||||||| ||||| ||||||||||||||
Db 30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1

RESULT 2
AAA86892
ID AAA86892 standard; DNA; 30 BP.
XX
AC AAA86892;
XX
DT 15-JAN-2001 (first entry)
XX
DE Probe to Campylobacter jejuni.
XX
KW Detection; nucleic acid hybrid; depolymerisation; analysis; SNP;
KW single nucleotide polymorphism; identification; viral load; probe;
KW genotyping; medical marker diagnostic; primer; target; mutation;
KW genetic disease; ss.
XX
OS Campylobacter jejuni.
XX
PN WO200049180-A1.
XX
PD 24-AUG-2000.
XX
PF 18-FEB-2000; 2000WO-US04242.
XX
PR 18-FEB-1999; 99US-0252436.
PR 21-JUL-1999; 99US-0358972.
PR 25-AUG-1999; 99US-0383316.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Shultz JW, Lewis MK, Leippe D, Mandrekar M, Kephart D, Rhodes RB;
PI Andrews CA, Hartnett JR, Gu T, Olson RJ, Wood KV, Welch R;
XX
DR WPI; 2000-565377/52.
XX
PT Determining presence or absence of a predetermined endogenous nucleic
PT acid sequence by using an enzyme that depolymerizes the 3' end of an
PT oligonucleotide probe hybridized to a target sequence to release
PT identifier nucleotides -
XX
PS Example; Page 321; 389pp; English.
XX
CC The present invention describes a method (M1) for determining the
CC presence or absence of a predetermined endogenous nucleic acid target
CC sequence (ENAT). The method comprises hybridising a probe having an
CC identifier nucleotide (IN) with ENAT which is treated with an enzyme
CC that depolymerises the 3' end of hybridised NA to release the INs.
CC M1 is used for determining the number of known sequence repeats present
CC in a nucleic acid target sequence in a nucleic acid sample. The method
CC is also useful for determining whether a nucleic acid target sequence in
CC a sample is an allele from a homozygous or heterozygous locus. The
CC method is also useful for detection of mutations, translocations and
CC SNPs in nucleic acids (including those associated with genetic disease),
CC determination of viral load, species identification, sample
CC contamination, and analysis of forensic samples. AAA86791 to AAA87079
CC and AAB12817 represent sequence which are used in the exemplification of
CC the present invention.
CC N.B. There is a discrepancy between the SEQ ID NO: and sequences given
CC in the examples, and the SEQ ID NO: and sequences given in the sequence
CC listing from the present invention.
XX
SQ Sequence 30 BP; 16 A; 4 C; 5 G; 5 T; 0 other;

Query Match 2.7%; Score 26.8; DB 21; Length 30;
Best Local Similarity 93.3%; Pred. No. 1.3e+03;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 289 caagatggtcaaagcttaaaaacaagaact 318
||||||| ||||| ||||||||||||||
Db 1 caagatggacaaagtttaaaaacaagaact 30

RESULT 3
AAA93188/c
ID AAA93188 standard; DNA; 30 BP.
XX
AC AAA93188;
XX
DT 11-JAN-2001 (first entry)
XX
DE Campylobacter jejuni interrogation probe 11451.
XX
KW Campylobacter jejuni; nucleic acid detection; genomic typing;
KW mutation detection; viral load determination; species identification;
KW forensic analysis; probe; ss.
XX
OS Campylobacter jejuni.
XX
PN WO200049179-A1.
XX
PD 24-AUG-2000.
XX
PF 18-FEB-2000; 2000WO-US04176.
XX
PR 18-FEB-1999; 99US-0252436.
PR 21-JUL-1999; 99US-0358972.
PR 27-SEP-1999; 99US-0406147.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Shultz JW, Lewis MK, Leippe D, Mandrekar M, Kephart D, Rhodes RB;
PI Andrews CA, Hartnett JR, Gu T, Olson RJ, Wood KV, Welch R;
XX
DR WPI; 2000-549282/50.
XX
PT Detecting the presence of predetermined exogenous nucleic acid target
PT sequence useful for e.g. genotyping, comprises depolymerizing the 3'
PT end of an oligonucleotide probe hybridized to a nucleic acid target

PT sequence -
XX
PS Claim 47; Page 187; 230pp; English.
XX
CC The present sequence is an interrogation probe which was used to detect a
CC segment of the genome of Campylobacter jejuni. This was performed as part
CC of a method for determining the presence of a known exogenous nucleic
CC acid target sequence in a nucleic acid sample. The method comprises
CC admixing a treated sample with a depolymerising enzyme which releases one
CC or more nucleotides from the 3'-end of a hybridised nucleic acid probe.
CC The method is used for assaying nucleic acids for a particular native or
CC mutant sequence, and for genomic typing. It is useful for detecting
CC mutations, translocations, and single nucleotide polymorphisms,
CC determination of viral load, species identification, detection of sample
CC contamination, and analysis of forensic samples. Compared with previous
CC methods of detecting nucleic acid hybrids, the new method has higher
CC sensitivity without the need for radiochemicals or electrophoresis. It is
CC quantitative, highly reproducible and can be automated. The method can
CC reliably detect as few as 10 copies of a virus in a sample, and is
CC capable of providing multiple analyses in a single assay (multiplex
CC assay).
XX
SQ Sequence 30 BP; 5 A; 5 C; 4 G; 16 T; 0 other;

Query Match 2.7%; Score 26.8; DB 21; Length 30;
Best Local Similarity 93.3%; Pred. No. 1.3e+03;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 289 caagatggtcaaagcttaaaaacaagaact 318
||||||||| ||||| ||||||||||||||
Db 30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1

RESULT 4
AAA93190
ID AAA93190 standard; DNA; 30 BP.
XX
AC AAA93190;
XX
DT 11-JAN-2001 (first entry)
XX
DE Campylobacter jejuni interrogation probe 11450.
XX
KW Campylobacter jejuni; nucleic acid detection; genomic typing;
KW mutation detection; viral load determination; species identification;
KW forensic analysis; probe; ss.
XX
OS Campylobacter jejuni.
XX
PN WO200049179-A1.
XX
PD 24-AUG-2000.
XX
PF 18-FEB-2000; 2000WO-US04176.
XX
PR 18-FEB-1999; 99US-0252436.
PR 21-JUL-1999; 99US-0358972.
PR 27-SEP-1999; 99US-0406147.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Shultz JW, Lewis MK, Leippe D, Mandrekar M, Kephart D, Rhodes RB;
PI Andrews CA, Hartnett JR, Gu T, Olson RJ, Wood KV, Welch R;
XX
DR WPI; 2000-549282/50.
XX
PT Detecting the presence of predetermined exogenous nucleic acid target
PT sequence useful for e.g. genotyping, comprises depolymerizing the 3'
PT end of an oligonucleotide probe hybridized to a nucleic acid target
PT sequence -
XX
PS Claim 47; Page 187; 230pp; English.
XX
CC The present sequence is an interrogation probe which was used to detect a
CC segment of the genome of Campylobacter jejuni. This was performed as part
CC of a method for determining the presence of a known exogenous nucleic
CC acid target sequence in a nucleic acid sample. The method comprises
CC admixing a treated sample with a depolymerising enzyme which releases one
CC or more nucleotides from the 3'-end of a hybridised nucleic acid probe.
CC The method is used for assaying nucleic acids for a particular native or
CC mutant sequence, and for genomic typing. It is useful for detecting
CC mutations, translocations, and single nucleotide polymorphisms,
CC determination of viral load, species identification, detection of sample
CC contamination, and analysis of forensic samples. Compared with previous
CC methods of detecting nucleic acid hybrids, the new method has higher
CC sensitivity without the need for radiochemicals or electrophoresis. It is
CC quantitative, highly reproducible and can be automated. The method can
CC reliably detect as few as 10 copies of a virus in a sample, and is
CC capable of providing multiple analyses in a single assay (multiplex
CC assay).
XX
SQ Sequence 30 BP; 16 A; 4 C; 5 G; 5 T; 0 other;

Query Match 2.7%; Score 26.8; DB 21; Length 30;
Best Local Similarity 93.3%; Pred. No. 1.3e+03;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 289 caagatggtcaaagcttaaaaacaagaact 318
||||||||| ||||| ||||||||||||||
Db 1 caagatggacaaagtttaaaaacaagaact 30

RESULT 5
AAX52169/c
ID AAX52169 standard; DNA; 50 BP.
XX
AC AAX52169;
XX
DT 18-JUN-1999 (first entry)
XX
DE Synthetic plasmid synlux4 construction oligonucleotide R53.
XX
KW DNA plasmid; lux A; lux B; Vibrio fisherii; luciferase; promoter;
KW tn9 kanamycin/neomycin phosphotransferase; DNA synthesis;
KW replication competent double-stranded polynucleotide; ss.
XX
OS Synthetic.
XX
PN WO9914318-A1.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US19312.
XX
PR 16-SEP-1997; 97US-0059017.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Evans GA;
XX
DR WPI; 1999-244029/20.
XX
PT Synthesis of replication competent double-stranded polynucleotides
XX
PS Example 4; Fig 5E; 135pp; English.
XX
CC AAX52021-212 represent oligonculeotide primers that were used to
CC construct a synthetic DNA plasmid sequence synlux4, to demonstrate the
CC method of the invention. Within the synlux4 sequence are included the
CC sequences of lux A, lux B, the A and B components of the Vibrio fisherii
CC luciferase sequence, positions of pUC19 including the origin of
CC replication and replication stability sequences, and the promoter and
CC coding sequence for tn9 kanamycin/neomycin phosphotransferase. The
CC specification describes a method for the synthesis of replication

```
CC   competent double-stranded polynucleotides. The method comprises
CC   generating a first set of oligonucleotides corresponding to the plus
CC   strand and a second set corresponding to the minus strand and
CC   annealing. The method can be used for preparing polynucleotides
CC   encoding sequences involved in a biochemical pathway. In particular,
CC   they can be used to produce polynucleotides encoding enzymes,
CC   e.g. hexokinase, phosphohexose isomerase, phosphofructokinase-1,
CC   aldolase, triose-phosphate isomerase, glyceraldehyde-3-phosphate
CC   dehydrogenase, phosphoglycerate kinase, phosphoglycerate mutase,
CC   enolase or pyruvate kinase. They can also be used for the preparation
CC   of viral particles, artificial genomes and artificial genetic systems.
XX
SQ   Sequence 50 BP; 19 A; 8 C; 4 G; 19 T; 0 other;


  Query Match             2.2%;  Score 21.8;  DB 20;  Length 50;
  Best Local Similarity   70.7%;  Pred. No. 2.3e+04;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    808 ttaatttcagctatcaatgctgtaaaagatacaactggtgt 848
          | ||| || | ||| || | ||||||||||  | || | |
Db     43 TCAATGTCCGATATTAAAGATGTAAAAGATATTATTGATAT 3


RESULT   6
AAV04223
ID   AAV04223 standard; DNA; 51 BP.
XX
AC   AAV04223;
XX
DT   22-JUN-1998  (first entry)
XX
DE   Human cardiac troponin I/troponin C 3' PCR primer.
XX
KW   Troponin I; troponin C; immunoassay; assay; analysis; human;
KW   cardiac muscle; skeletal muscle; injury; myocardial infarction;
KW   diagnosis; HcTnI; HcTnC; PCR; primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9739132-A1.
XX
PD   23-OCT-1997.
XX
PF   14-APR-1997;   97WO-US06147.
XX
PR   11-APR-1997;   97US-0833743.
PR   16-APR-1996;   96US-0015772.
XX
PA   (UYMI-) UNIV MIAMI.
XX
PI   Potter JD;
XX
DR   WPI; 1998-062676/06.
XX
PT   Immunoassay of mammalian troponin using stable standard for
PT   comparison - specifically acid-dialysed solution or its lyophilisate
PT   used for diagnosis of cardiac or skeletal muscle damage
XX
PS   Example 3; Page 34; 94pp; English.
XX
CC   This 3' PCR primer was used in the amplification of human cardiac
CC   troponin I (HcTnI) cDNA.  It is a complementary sequence encoding
CC   the C-terminal 8 amino acids of HcTnI followed by the N-terminal 8
CC   amino acids of human cardiac troponin C (HcTnC).  It was used with
CC   a vector-based 5' primer in the PCR amplification of HcTnI plasmid
CC   DNA.  HcTnC DNA was also amplified (see AAV04223), and the PCR
CC   products were used to construct a polynucleotide (see AAV04225)
CC   encoding a HcTnI/HcTnC fusion protein (see AAW41571).  The addition
CC   of the calcium binding protein HcTnC to HcTnI was made to provide
CC   more favourable solubility properties.  The fusion protein can be
CC   used as a standard in novel assays of mammalian troponin.
XX
SQ   Sequence 51 BP; 8 A; 17 C; 9 G; 17 T; 0 other;


  Query Match             2.1%;  Score 21.4;  DB 19;  Length 51;
  Best Local Similarity   71.8%;  Pred. No. 2.9e+04;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    505 agatttgaaacaggttcacaaagtttttcttcaggcact 543
          |||| | |  || | || |||| ||||||||  ||| ||
Db     12 agatgtcatccatgctctcaaacttttttcttgcggccct 50


RESULT   7
AAV04224/c
ID   AAV04224 standard; DNA; 58 BP.
XX
AC   AAV04224;
XX
DT   22-JUN-1998  (first entry)
XX
DE   Human cardiac troponin C 5' PCR primer.
XX
KW   Troponin C; troponin I; immunoassay; assay; analysis; human;
KW   cardiac muscle; skeletal muscle; injury; myocardial infarction;
KW   diagnosis; HcTnC; HcTnI; PCR; primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9739132-A1.
XX
PD   23-OCT-1997.
XX
PF   14-APR-1997;   97WO-US06147.
XX
PR   11-APR-1997;   97US-0833743.
PR   16-APR-1996;   96US-0015772.
XX
PA   (UYMI-) UNIV MIAMI.
XX
PI   Potter JD;
XX
DR   WPI; 1998-062676/06.
XX
PT   Immunoassay of mammalian troponin using stable standard for
PT   comparison - specifically acid-dialysed solution or its lyophilisate
PT   used for diagnosis of cardiac or skeletal muscle damage
XX
PS   Example 3; Page 34; 94pp; English.
XX
CC   This 5' PCR primer was used with a vector-based 3' primer in the
CC   amplification of human cardiac troponin C (HcTnC) DNA.  Human
CC   cardiac troponin I (HcTnI) DNA was also amplified (see AAV04223) and
CC   the PCR products were used to construct a polynucleotide (see
CC   AAV04225) encoding a HcTnI/HcTnC fusion protein (see AAW41571).  The
CC   addition of the calcium binding protein HcTnC to HcTnI was made to
CC   provide more favourable solubility properties.  The fusion protein
CC   can be used as a standard in novel assays of mammalian troponin.
XX
SQ   Sequence 58 BP; 18 A; 10 C; 20 G; 10 T; 0 other;


  Query Match             2.1%;  Score 21.4;  DB 19;  Length 58;
  Best Local Similarity   71.8%;  Pred. No. 3e+04;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    505 agatttgaaacaggttcacaaagtttttcttcaggcact 543
          |||| | |  || | || |||| ||||||||  ||| ||
Db     39 AGATGTCATCCATGCTCTCAAACTTTTTCTTGCGGCCCT 1


RESULT   8
```

```
AAA97417
ID   AAA97417 standard; DNA; 54 BP.
XX
AC   AAA97417;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Pea wild-type pra2 gene light-repressible promoter oligonucleotide, WT3.
XX
KW   GTP-binding protein pra2; pea; light-repressible promoter;
KW   photoinhibitory; expression cassette; transgenic plant;
KW   deterioration prevention; storage; ss.
XX
OS   Pisum sativum.
XX
PN   WO200055313-A1.
XX
PD   21-SEP-2000.
XX
PF   03-MAR-2000; 2000WO-JP01269.
XX
PR   12-MAR-1999;   99JP-0066551.
XX
PA   (SUNR ) SUNTORY LTD.
XX
PI   Sasaki Y,  Nagano Y,  Inaba T;
XX
DR   WPI; 2000-587526/55.
XX
PT   New DNA fragment or promoter for expressing a target gene, specifically
PT   under photoinhibitory conditions, and for transforming a plant cell or
PT   plant to improve quality and prevent deterioration during storage  -
XX
PS   Example 9; Page 19; 49pp; Japanese.
XX
CC   The invention relates to a light-repressible promoter (AAA97385), or
CC   active fragments thereof (AAA97383, AAA97384), from the gene encoding
CC   the pea GTP-binding protein pra2. The invention also relates to an
CC   expression cassette containing the pra2 promoter or its active
CC   fragments for the expression of a gene under photoinhibitory or dark
CC   conditions in a plant, and to transgenic plants, their descendants
CC   and plant tissues comprising the expression cassette. The expression
CC   cassette of the invention can be used to generate transgenic plants in
CC   which deterioration during storage in the dark is prevented. This is
CC   particularly useful for agricultural products. Sequences AAA97417-A97418
CC   represent oligonucleotides used in an exemplification of the invention
CC   to generate a wild-type pea pra2 promoter fragment.
XX
SQ   Sequence 54 BP; 18 A; 3 C; 12 G; 21 T; 0 other;


  Query Match             2.1%;  Score 21.2;  DB 21;  Length 54;
  Best Local Similarity  64.0%;  Pred. No. 3.3e+04;
  Matches   32;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    738 tcaggattttgctatcaatggggttgttataggtaaggttgattattcag 787
          | |||||||| |  | | || || | ||| ||  |  |  |||| | |||
Db      1 tgaggattttacagtaattgaggattttacagtaattgaggattttacag 50


RESULT   9
AAA27148
ID   AAA27148 standard; DNA; 27 BP.
XX
AC   AAA27148;
XX
DT   11-SEP-2000  (first entry)
XX
DE   Campylobacter coli flaA gene primer flaA-11.
XX
KW   Flagellin; flaA; diarrhoea; Guillain-Barre syndrome;
KW   vaccine; GBS; PCR primer; ss.
XX
OS   Campylobacter coli.
XX
PN   WO200027205-A1.
XX
PD   18-MAY-2000.
XX
PF   12-NOV-1999;   99WO-US27195.
XX
PR   12-NOV-1998;   98US-0108114.
XX
PA   (USAT ) US SEC.
XX
PI   Guerry P,  Trust T,  Burg E,  Lee L;
XX
DR   WPI; 2000-376214/32.
XX
PT   Campylobacter FlaA protein and coding sequence, useful in reducing
PT   Campylobacter intestinal colonization  -
XX
PS   Disclosure; Page 7; 43pp; English.
XX
CC   The flaA gene encodes the major flagellin subunit of the Campylobacter
CC   coli flagellar filament. Part of the FlaA polypeptide may be fused with
CC   the maltose binding protein of Escherichia coli  to make a recombinant
CC   protein. When this protein is introduced into a host an immunological
CC   response is triggered. Therefore the recombinant protein may be used as
CC   a vaccine to protect against C. coli intestinal colonisation and the
CC   diarrhoea it causes. This vaccine system is useful as it can
CC   prevent the development of Guillain-Barre syndrome (GBS) which is seen
CC   with whole cell Campylobacter vaccines. The present sequence is the
CC   flaA-11 PCR primer that was used to amplify part of the flaA gene.
XX
SQ   Sequence 27 BP; 12 A; 6 C; 3 G; 6 T; 0 other;


  Query Match             2.1%;  Score 21;  DB 21;  Length 27;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+04;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 attaacacaaatgttgcagca 21
          |||||||||||||||||||||
Db      7 attaacacaaatgttgcagca 27


RESULT  10
AAA27149/c
ID   AAA27149 standard; DNA; 33 BP.
XX
AC   AAA27149;
XX
DT   11-SEP-2000  (first entry)
XX
DE   Campylobacter coli flaA gene primer flaA-2.
XX
KW   Flagellin; flaA; diarrhoea; Guillain-Barre syndrome;
KW   vaccine; GBS; PCR primer; ss.
XX
OS   Campylobacter coli.
XX
PN   WO200027205-A1.
XX
PD   18-MAY-2000.
XX
PF   12-NOV-1999;   99WO-US27195.
XX
PR   12-NOV-1998;   98US-0108114.
XX
PA   (USAT ) US SEC.
XX
PI   Guerry P,  Trust T,  Burg E,  Lee L;
XX
DR   WPI; 2000-376214/32.
XX
```

PT Campylobacter FlaA protein and coding sequence, useful in reducing
PT Campylobacter intestinal colonization -
XX
PS Disclosure; Page 7; 43pp; English.
XX
CC The flaA gene encodes the major flagellin subunit of the Campylobacter
CC coli flagellar filament. Part of the FlaA polypeptide may be fused with
CC the maltose binding protein of Escherichia coli to make a recombinant
CC protein. When this protein is introduced into a host an immunological
CC response is triggered. Therefore the recombinant protein may be used as
CC a vaccine to protect against C. coli intestinal colonisation and the
CC diarrhoea it causes. This vaccine system is useful as it can
CC prevent the development of Guillain-Barre syndrome (GBS) which is seen
CC with whole cell Campylobacter vaccines. The present sequence is the
CC flaA-2 PCR primer that was used to amplify part of the flaA gene.
XX
SQ Sequence 33 BP; 10 A; 8 C; 1 G; 14 T; 0 other;

Query Match 2.1%; Score 21; DB 21; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 979 gttaaaaatgatggtagagat 999
||||||||||||||||||||
Db 33 GTTAAAAATGATGGTAGAGAT 13

RESULT 11
AAT60508/c
ID AAT60508 standard; DNA; 24 BP.
XX
AC AAT60508;
XX
DT 10-JUN-1997 (first entry)
XX
DE Primer CFO4R.4.
XX
KW PCR; polymerase chain reaction; amplify; infection; forensic science;
KW infectious pathogen; genetic disorder; genetic variance; primer; ss.
XX
OS Synthetic.
XX
PN US5612473-A.
XX
PD 18-MAR-1997.
XX
PF 16-JAN-1996; 96US-0587209.
XX
PR 16-JAN-1996; 96US-0587209.
XX
PA (GULL-) GULL LAB.
XX
PI Coombs J, Glass MJ, Malmstrom SL, Wu L;
XX
DR WPI; 1997-192163/17.
XX
PT Processing samples for amplification of nucleic acid target
PT sequences - using extraction buffer containing at least one
PT detergent and a salt composition of greater that 1 molar
PT concentration
XX
PS Example 3; Column 17; 21pp; English.
XX
CC AAT60503-T60514 represent amplification primers for DNA sequences
CC present in a sample processed by the method of the invention. The
CC processing method of the invention comprises obtaining a sample of
CC material potentially containing the target nucleic acid sequences, and
CC mixing the sample with an external buffer solution. The buffer solution
CC comprises two detergents, and at least one salt composition present in a
CC greater than 1 M concentration. The mixture is then centrifuged to obtain
CC a supernatant portion, which is then heated before being recentrifuged
CC to precipitate the proteins, and obtaining a second supernatant portion,
CC from which nucleic acids are precipitated. The isolated nucleic acids
CC are then dissolved. The method provides a rapid means of preparing a
CC sample for amplification so that multiple analytes can be detected and
CC differentiated within a relatively short time period (typically less
CC than 5 hours with the novel pre-processing step taking less than 5
CC minutes). Typical applications of nucleic acid amplification include
CC detection of infections in patients, foodstuffs and for
CC diagnostic/forensic or quality control purposes, to discriminate between
CC multiple potential infectious pathogens, to diagnose genetic disorders or
CC to identify genetic variances.
XX
SQ Sequence 24 BP; 6 A; 5 C; 5 G; 8 T; 0 other;

Query Match 2.1%; Score 20.8; DB 18; Length 24;
Best Local Similarity 91.7%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 91 ggtcttagaatcaactccgcagca 114
|||||||||| ||||| ||||||
Db 24 GGTCTTAGAATTAACTCAGCAGCA 1

RESULT 12
AAV31446/c
ID AAV31446 standard; DNA; 24 BP.
XX
AC AAV31446;
XX
DT 11-AUG-1998 (first entry)
XX
DE Campylobacter nucleic acid sequence amplifying primer CFO4R.
XX
KW Salmonella; microorganism; detection; multiple analyte; PCR primer;
KW Yersinia; Escherichia coli; Campylobacter; ss.
XX
OS Synthetic.
OS Campylobacter sp.
XX
PN US5756701-A.
XX
PD 26-MAY-1998.
XX
PF 06-AUG-1996; 96US-0692725.
XX
PR 16-JAN-1996; 96US-0587209.
PR 06-AUG-1996; 96US-0692725.
XX
PA (GULL-) GULL LAB INC.
XX
PI Coombs J, Glass MJ, Malmstrom SL, Wu L;
XX
DR WPI; 1998-321634/28.
XX
PT Nucleic acid probes and primers - for detecting Salmonella, Yersinia
PT or E. coli
XX
PS Claim 5; Column 17; 21pp; English.
XX
CC This primer is used for the PCR amplification of Campylobacter nucleic
CC acid sequences. The invention provides nucleic acid probes and primers
CC for detecting Salmonella, Yersinia or E. coli. It provides methods and
CC apparatus for detecting and discriminating multiple analytes within a
CC test sample. The methods are simple, user-friendly, cost effective and
CC fast. The methods and the probes and primer sequences are used for
CC detecting the corresponding microorganisms in clinical samples.
XX
SQ Sequence 24 BP; 6 A; 5 C; 5 G; 8 T; 0 other;

Query Match 2.1%; Score 20.8; DB 19; Length 24;
Best Local Similarity 91.7%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1
```

RESULT 13
AAV25942/c
ID AAV25942 standard; DNA; 24 BP.
XX
AC AAV25942;
XX
DT 15-JUL-1998 (first entry)
XX
DE Oligonucleotide PCR primer CFO4R gene.
XX
KW Sequence-specific; probe; enterohaemorrhagic; Escherichia coli;
KW Salmonella; Campylobacter; Shigella; Yersinia; beta-globin;
KW gastroenteritis; PCR primer; ss.
XX
OS Synthetic.
OS Campylobacter sp.
XX
PN US5753444-A.
XX
PD 19-MAY-1998.
XX
PF 07-AUG-1996; 96US-0689235.
XX
PR 16-JAN-1996; 96US-0587209.
PR 07-AUG-1996; 96US-0689235.
XX
PA (GULL-) GULL LAB INC.
XX
PI Coombs J, Glass MJ, Malmstrom SL, Wu L;
XX
DR WPI; 1998-311393/27.
XX
PT Distinguishing between similar nucleic acid samples - using
PT sequence-specific probes e.g. between enterohaemorrhagic and normal
PT Escherichia coli
XX
PS Example 3; Column 17; 21pp; English.
XX
CC The present sequence represents a PCR primer used in an example of the
CC present invention. The present invention describes a method for
CC detecting mismatches between first and second nucleic acid sequences
CC having at least one base difference. The method comprises: (a) obtaining
CC at least one labelled probe consisting of an oligonucleotide sequence
CC spanning the location of at least one base difference between the first
CC and second sequences, where the oligonucleotide sequence contains at
CC least one neutral base molecule in a position other than the position of
CC the base difference(s) but is otherwise exactly complementary to the
CC first sequence, so that the probe hybridises more weakly with the second
CC sequence than with the first sequence; (b) mixing the probe(s) with the
CC first and second sequences under hybridisation conditions; (c)
CC dissociating any probe/second sequence hybrids; and (d) detecting
CC probe/first sequence hybrids. The method can be used to distinguish
CC between similar DNA/RNA sequences in a sample, especially to distinguish
CC enterohaemorrhagic E. coli O157:H7 from other E. coli strains e.g. in
CC stool samples from people suffering from gastroenteritis, caused
CC specifically by enterohaemorrhagic E.coli. Use of the method shortens
CC the time between sample preparation to obtaining results, than has been
CC possible with previous similar procedures.
XX
SQ Sequence 24 BP; 6 A; 5 C; 5 G; 8 T; 0 other;

Query Match 2.1%; Score 20.8; DB 19; Length 24;
Best Local Similarity 91.7%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1
```

RESULT 14
AAV20847/c
ID AAV20847 standard; DNA; 24 BP.
XX
AC AAV20847;
XX
DT 01-JUL-1998 (first entry)
XX
DE Campylobacter CFO4R gene PCR primer.
XX
KW Escherichia coli strain O157:H7; detection; microorganism; infection;
KW enterohaemorrhagic; PCR primer; ss.
XX
OS Synthetic.
OS Campylobacter sp.
XX
PN US5738995-A.
XX
PD 14-APR-1998.
XX
PF 07-AUG-1996; 96US-0689236.
XX
PR 16-JAN-1996; 96US-0587209.
PR 07-AUG-1996; 96US-0689236.
XX
PA (GULL-) GULL LAB INC.
XX
PI Coombs J, Glass MJ, Malmstrom SL, Wu L;
XX
DR WPI; 1998-260031/23.
XX
PT Probes for detecting Escherichia coli strain O157:H7 - useful for
PT diagnosis of enterohaemorrhagic Escherichia coli infection(s)
XX
PS Example 3; Column 17; 21pp; English.
XX
CC The present sequence represents a PCR primer used in an example of the
CC present invention. The present invention describes probes used in the
CC detection of Escherichia coli strain O157:H7 in a sample. The method of
CC detection comprises: (a) obtaining at least 1 probe specifically given
CC in the specification, labelled with a label that permits probe detection
CC when hybridised to a complementary nucleic acid sequence which is
CC specific for a nucleic acid sequence of the microorganism; (b)
CC hybridising the probes and the sample, and (c) detecting hybrids
CC comprising the probes and the nucleic acid sequences. The method and
CC probes may be used for diagnosis of enterohaemorrhagic E. coli
CC infections. The methods and the materials permit the detection and
CC discrimination of multiple analytes.
XX
SQ Sequence 24 BP; 6 A; 5 C; 5 G; 8 T; 0 other;

Query Match 2.1%; Score 20.8; DB 19; Length 24;
Best Local Similarity 91.7%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1
```

RESULT 15
AAF16711
ID AAF16711 standard; DNA; 44 BP.
XX
AC AAF16711;
XX
DT 09-MAR-2001 (first entry)
XX

DE dGMP-specific aptamer clone #19.
XX
KW Polymeric biomolecule; aptamer; ss.
XX
OS Synthetic.
XX
PN WO200071755-A2.
XX
PD 30-NOV-2000.
XX
PF 25-MAY-2000; 2000WO-US14401.
XX
PR 25-MAY-1999; 99US-0135863.
XX
PA (PRAE-) PRAELUX INC.
XX
PI Kwagh J, Macklin JJ, Mitsis PG, Ulmer KM;
XX
DR WPI; 2001-016410/02.
XX
PT Sequencing a polymeric biomolecule, such as a polynucleotide,
PT polysaccharide or polypeptide, comprises separating a terminal monomer
PT from the polymeric biomolecule and identifying the separated terminal
PT monomer using an aptamer -
XX
PS Claim 37; Fig 15; 123pp; English.
XX
CC The present invention relates to a new method for sequencing a
CC polymeric biomolecule. The method involves separating a terminal
CC monomer from the polymeric biomolecule and identifying the separated
CC terminal monomer using an aptamer. The method is useful for
CC sequencing or structurally characterizing a polymeric biomolecule
CC such as a polynucleotide, a polysaccharide or a polypeptide. The
CC method is also useful for developing aptamers.
XX
SQ Sequence 44 BP; 9 A; 6 C; 19 G; 10 T; 0 other;

Query Match 2.1%; Score 20.8; DB 22; Length 44;
Best Local Similarity 70.0%; Pred. No. 3.9e+04;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 747 tgctatcaatggggttgttataggtaaggttgattattca 786
|| | || |||||| | ||| |||| ||||| | ||| .
Db 4 tgacaccactggggtgggtatgggtagggttgtggaatca 43

Search completed: April 17, 2002, 02:17:16
Job time: 3512 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 17, 2002, 01:23:59 ; Search time 86.97 Seconds
(without alignments)
2601.489 Million cell updates/sec

Title: US-09-439-311-1
Perfect score: 999
Sequence: 1 attaacacaaatgttgcagc.....ttaaaaatgatggtagagat 999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 515962

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	26.8	2.7	30	4	US-09-358-972-102	Sequence 102, App
	2	26.8	2.7	30	4	US-09-358-972-103	Sequence 103, App
c	3	26.8	2.7	30	4	US-09-406-147-32	Sequence 32, Appl
	4	26.8	2.7	30	4	US-09-406-147-34	Sequence 34, Appl
c	5	20.8	2.1	24	1	US-08-587-209-6	Sequence 6, Appli
c	6	20.8	2.1	24	1	US-08-689-236-6	Sequence 6, Appli
c	7	20.8	2.1	24	1	US-08-689-235-6	Sequence 6, Appli
c	8	20.8	2.1	24	1	US-08-692-725-6	Sequence 6, Appli
c	9	20.8	2.1	24	2	US-08-692-726-6	Sequence 6, Appli
c	10	20.8	2.1	58	2	US-08-431-527A-7	Sequence 7, Appli
c	11	20.8	2.1	58	4	US-09-214-278-30	Sequence 30, Appl
c	12	20.6	2.1	52	3	US-08-886-967-3	Sequence 3, Appli
c	13	20.6	2.1	52	4	US-09-306-949-3	Sequence 3, Appli
c	14	20	2.0	39	3	US-08-874-825-118	Sequence 118, App
c	15	20	2.0	50	2	US-08-450-905B-7	Sequence 7, Appli
c	16	20	2.0	50	3	US-07-982-759F-7	Sequence 7, Appli
	17	20	2.0	57	1	US-08-141-892A-4	Sequence 4, Appli
	18	20	2.0	57	2	US-08-583-447A-4	Sequence 4, Appli
	19	20	2.0	57	2	US-08-467-920-4	Sequence 4, Appli
	20	20	2.0	57	3	US-08-635-930-4	Sequence 4, Appli
	21	20	2.0	57	3	US-09-193-997-4	Sequence 4, Appli
	22	20	2.0	57	4	US-09-138-237A-4	Sequence 4, Appli
	23	20	2.0	60	1	US-08-478-370-4	Sequence 4, Appli
c	24	19.6	2.0	43	1	US-07-959-946-12	Sequence 12, Appl
c	25	19.6	2.0	43	1	US-08-333-577-12	Sequence 12, Appl
c	26	19.6	2.0	43	5	PCT-US92-08634-12	Sequence 12, Appl
c	27	19.6	2.0	58	1	US-08-105-483-174	Sequence 174, App
c	28	19.6	2.0	58	1	US-08-709-209-174	Sequence 174, App
c	29	19.6	2.0	58	1	US-08-303-275-62	Sequence 62, Appl
c	30	19.6	2.0	58	1	US-08-458-101-174	Sequence 174, App
	31	19.6	2.0	60	1	US-07-670-296-19	Sequence 19, Appl
	32	19.6	2.0	60	1	US-08-093-781-20	Sequence 20, Appl
c	33	19.6	2.0	60	3	US-08-963-602-2	Sequence 2, Appli
c	34	19.4	1.9	37	2	US-08-403-853-8	Sequence 8, Appli
	35	19.4	1.9	60	1	US-08-487-890A-127	Sequence 127, App
	36	19.4	1.9	60	2	US-08-478-435-127	Sequence 127, App
	37	19.4	1.9	60	2	US-08-337-483-127	Sequence 127, App
	38	19.4	1.9	60	2	US-08-478-373-127	Sequence 127, App
	39	19.4	1.9	60	3	US-08-474-671-127	Sequence 127, App
	40	19.4	1.9	60	3	US-08-483-577A-127	Sequence 127, App
	41	19.4	1.9	60	4	US-08-897-438-127	Sequence 127, App
c	42	19.2	1.9	58	2	US-08-431-527A-6	Sequence 6, Appli
	43	19.2	1.9	60	1	US-07-609-716-72	Sequence 72, Appl
	44	19.2	1.9	60	3	US-08-475-411A-72	Sequence 72, Appl
	45	19.2	1.9	60	4	US-08-478-029A-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-09-358-972-102/c
; Sequence 102, Application US/09358972
; Patent No. 6235480
; GENERAL INFORMATION:
; APPLICANT: Shultz, John W
; APPLICANT: Lewis, Martin K.
; APPLICANT: Lieppe, Donna
; APPLICANT: Mandrekar, Michelle
; APPLICANT: Kephart, Daniel
; APPLICANT: Rhodes, Richard B.
; APPLICANT: Andrews, Christine A.
; APPLICANT: Hartnett, James R.
; APPLICANT: Gu, Trent
; APPLICANT: Olson, Ryan J.
; APPLICANT: Wood, Keith W.
; APPLICANT: Welch, Roy
; TITLE OF INVENTION: Nucleic Acid Detection
; FILE REFERENCE: Pro-103 6868/75528
; CURRENT APPLICATION NUMBER: US/09/358,972
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: 09/252,436
; EARLIER FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: 09/042,287
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: probe to Campylobacter jejuni
US-09-358-972-102

Query Match 2.7%; Score 26.8; DB 4; Length 30;
Best Local Similarity 93.3%; Pred. No. 82;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    289 caagatggtcaaagcttaaaaacaagaact 318
          ||||||| ||||| ||||||||||||||
Db     30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1
```

RESULT 2
US-09-358-972-103
; Sequence 103, Application US/09358972
; Patent No. 6235480
; GENERAL INFORMATION:

```
;  APPLICANT: Shultz, John W
;  APPLICANT: Lewis, Martin K.
;  APPLICANT: Lieppe, Donna
;  APPLICANT: Mandrekar, Michelle
;  APPLICANT: Kephart, Daniel
;  APPLICANT: Rhodes, Richard B.
;  APPLICANT: Andrews, Christine A.
;  APPLICANT: Hartnett, James R.
;  APPLICANT: Gu, Trent
;  APPLICANT: Olson, Ryan J.
;  APPLICANT: Wood, Keith W.
;  APPLICANT: Welch, Roy
;  TITLE OF INVENTION: Nucleic Acid Detection
;  FILE REFERENCE: Pro-103 6868/75528
;  CURRENT APPLICATION NUMBER: US/09/358,972
;  CURRENT FILING DATE: 1999-07-22
;  EARLIER APPLICATION NUMBER: 09/252,436
;  EARLIER FILING DATE: 1999-02-18
;  EARLIER APPLICATION NUMBER: 09/042,287
;  EARLIER FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 290
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
;    LENGTH: 30
;    TYPE: DNA
;    ORGANISM: Campylobacter jejuni
;    FEATURE:
;    OTHER INFORMATION: probe to Campylobacter jejuni
US-09-358-972-103


  Query Match             2.7%;  Score 26.8;  DB 4;  Length 30;
  Best Local Similarity   93.3%;  Pred. No. 82;
  Matches   28;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    289 caagatggtcaaagcttaaaaacaagaact 318
          |||||||| ||||| ||||||||||||||
Db      1 caagatggacaaagtttaaaaacaagaact 30


RESULT   3
US-09-406-147-32/c
; Sequence 32, Application US/09406147
; Patent No. 6270974
; GENERAL INFORMATION:
;  APPLICANT: Shultz, John W
;  APPLICANT: Lewis, Martin K
;  APPLICANT: Leippe, Donna
;  APPLICANT: Mandrekar, Michelle
;  APPLICANT: Kephart, Daniel
;  APPLICANT: Rhodes, Richard B
;  APPLICANT: Andrews, Christine A
;  APPLICANT: Hartnett, James R
;  APPLICANT: Gu, Trent
;  APPLICANT: Wood, Keith V
;  APPLICANT: Welch, Roy
;  TITLE OF INVENTION: EXOGENOUS NUCLEIC ACID DETECTION
;  FILE REFERENCE: EXOGENOUS NUCLEIC ACID DETECTION
;  CURRENT APPLICATION NUMBER: US/09/406,147
;  CURRENT FILING DATE: 1999-09-27
;  EARLIER APPLICATION NUMBER: 09/252,436
;  EARLIER FILING DATE: 1999-02-18
;  EARLIER APPLICATION NUMBER: 09/042,287
;  EARLIER FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
;    LENGTH: 30
;    TYPE: DNA
;    ORGANISM: Campylobacter jejuni
US-09-406-147-32


  Query Match             2.7%;  Score 26.8;  DB 4;  Length 30;
  Best Local Similarity   93.3%;  Pred. No. 82;
  Matches   28;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    289 caagatggtcaaagcttaaaaacaagaact 318
          |||||||| ||||| ||||||||||||||
Db     30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1


RESULT   4
US-09-406-147-34
; Sequence 34, Application US/09406147
; Patent No. 6270974
; GENERAL INFORMATION:
;  APPLICANT: Shultz, John W
;  APPLICANT: Lewis, Martin K
;  APPLICANT: Leippe, Donna
;  APPLICANT: Mandrekar, Michelle
;  APPLICANT: Kephart, Daniel
;  APPLICANT: Rhodes, Richard B
;  APPLICANT: Andrews, Christine A
;  APPLICANT: Hartnett, James R
;  APPLICANT: Gu, Trent
;  APPLICANT: Wood, Keith V
;  APPLICANT: Welch, Roy
;  TITLE OF INVENTION: EXOGENOUS NUCLEIC ACID DETECTION
;  FILE REFERENCE: EXOGENOUS NUCLEIC ACID DETECTION
;  CURRENT APPLICATION NUMBER: US/09/406,147
;  CURRENT FILING DATE: 1999-09-27
;  EARLIER APPLICATION NUMBER: 09/252,436
;  EARLIER FILING DATE: 1999-02-18
;  EARLIER APPLICATION NUMBER: 09/042,287
;  EARLIER FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
;    LENGTH: 30
;    TYPE: DNA
;    ORGANISM: Campylobacter jejuni
US-09-406-147-34


  Query Match             2.7%;  Score 26.8;  DB 4;  Length 30;
  Best Local Similarity   93.3%;  Pred. No. 82;
  Matches   28;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    289 caagatggtcaaagcttaaaaacaagaact 318
          |||||||| ||||| ||||||||||||||
Db      1 caagatggacaaagtttaaaaacaagaact 30


RESULT   5
US-08-587-209-6/c
; Sequence 6, Application US/08587209
; Patent No. 5612473
;  GENERAL INFORMATION:
;    APPLICANT:  Wu, Linxian
;    APPLICANT:  Coombs, Jana
;    APPLICANT:  Malmstrom, Sharon L.
;    APPLICANT:  Glass, Michael J.
;    TITLE OF INVENTION:  Methods and Apparatus for Preparing, Amplifying,
;    TITLE OF INVENTION:  and Discriminating Multiple Analytes
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David O. Seeley, Esq.
;      ADDRESSEE:  Workman, Nydegger & Seeley
;      STREET:  1000 Eagle Gate Tower
;      STREET:  60 East South Temple
;      CITY:  Salt Lake City
;      STATE:  Utah
;      COUNTRY:  USA
;      ZIP:  84111
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Diskette, 3.50 inch,
;       MEDIUM TYPE:  1.44 Mb storage
;       COMPUTER:  IBM compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 6.0a for WINDOWS
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/587,209
;       FILING DATE:  16-JAN-1996
;       CLASSIFICATION:  435
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-587-209-6


  Query Match             2.1%;  Score 20.8;  DB 1;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 3e+03;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   6
US-08-689-236-6/c
; Sequence 6, Application US/08689236
; Patent No. 5738995
;  GENERAL INFORMATION:
;    APPLICANT:  Wu, Linxian
;    APPLICANT:  Coombs, Jana
;    APPLICANT:  Malmstrom, Sharon L.
;    APPLICANT:  Glass, Michael J.
;    TITLE OF INVENTION:  Methods and Apparatus for
;    TITLE OF INVENTION:  Preparing, Amplifying, and Discriminating Multiple Analytes
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David O. Seeley, Esq.
;      ADDRESSEE:  Workman, Nydegger & Seeley
;      STREET:  1000 Eagle Gate Tower
;      STREET:  60 East South Temple
;      CITY:  Salt Lake City
;      STATE:  Utah
;      COUNTRY:  USA
;      ZIP:  84111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch,
;      MEDIUM TYPE:  1.44 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0a for WINDOWS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/689,236
;      FILING DATE:  16-JAN-1996
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-689-236-6


  Query Match             2.1%;  Score 20.8;  DB 1;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 3e+03;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   7
US-08-689-235-6/c
; Sequence 6, Application US/08689235
; Patent No. 5753444
;  GENERAL INFORMATION:
;    APPLICANT:  Wu, Linxian
;    APPLICANT:  Coombs, Jana
;    APPLICANT:  Malmstrom, Sharon L.
;    APPLICANT:  Glass, Michael J.
;    TITLE OF INVENTION:  Methods and Apparatus for
;    TITLE OF INVENTION:  Preparing, Amplifying, and Discriminating Multiple Analytes
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David O. Seeley, Esq.
;      ADDRESSEE:  Workman, Nydegger & Seeley
;      STREET:  1000 Eagle Gate Tower
;      STREET:  60 East South Temple
;      CITY:  Salt Lake City
;      STATE:  Utah
;      COUNTRY:  USA
;      ZIP:  84111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch,
;      MEDIUM TYPE:  1.44 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0a for WINDOWS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/689,235
;      FILING DATE:  16-JAN-1996
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-689-235-6


  Query Match             2.1%;  Score 20.8;  DB 1;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 3e+03;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           ||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   8
US-08-692-725-6/c
; Sequence 6, Application US/08692725
; Patent No. 5756701
;  GENERAL INFORMATION:
;    APPLICANT:  Wu, Linxian
;    APPLICANT:  Coombs, Jana
;    APPLICANT:  Malmstrom, Sharon L.
;    APPLICANT:  Glass, Michael J.
;    TITLE OF INVENTION:  Methods and Apparatus for
;    TITLE OF INVENTION:  Preparing, Amplifying, and Discriminating Multiple Analytes
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David O. Seeley, Esq.
;      ADDRESSEE:  Workman, Nydegger & Seeley
;      STREET:  1000 Eagle Gate Tower
;      STREET:  60 East South Temple
;      CITY:  Salt Lake City
```

```
;       STATE:  Utah
;       COUNTRY:  USA
;       ZIP:  84111
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch,
;       MEDIUM TYPE:  1.44 Mb storage
;       COMPUTER:  IBM compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 6.0a for WINDOWS
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/692,725
;       FILING DATE:  16-JAN-1996
;       CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-692-725-6


  Query Match             2.1%;  Score 20.8;  DB 1;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 3e+03;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           |||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT   9
US-08-692-726-6/c
; Sequence 6, Application US/08692726
; Patent No. 5846783
;  GENERAL INFORMATION:
;     APPLICANT:  Wu, Linxian
;     APPLICANT:  Coombs, Jana
;     APPLICANT:  Malmstrom, Sharon L.
;     APPLICANT:  Glass, Michael J.
;     TITLE OF INVENTION:  Methods and Apparatus for Preparing,
;     TITLE OF INVENTION:  Amplifying, and Discriminating Multiple Analytes
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  David O. Seeley, Esq.
;       ADDRESSEE:  Workman, Nydegger & Seeley
;       STREET:  1000 Eagle Gate Tower
;       STREET:  60 East South Temple
;       CITY:  Salt Lake City
;       STATE:  Utah
;       COUNTRY:  USA
;       ZIP:  84111
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch,
;       MEDIUM TYPE:  1.44 Mb storage
;       COMPUTER:  IBM compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 6.0a for WINDOWS
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/692,726
;       FILING DATE:  06-AUG-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/587,209
;       FILING DATE:  16-JAN-1996
;  INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-692-726-6


  Query Match             2.1%;  Score 20.8;  DB 2;  Length 24;
  Best Local Similarity   91.7%;  Pred. No. 3e+03;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      91 ggtcttagaatcaactccgcagca 114
           |||||||||||| ||||| ||||||
Db      24 GGTCTTAGAATTAACTCAGCAGCA 1


RESULT  10
US-08-431-527A-7/c
; Sequence 7, Application US/08431527A
; Patent No. 5843650
;  GENERAL INFORMATION:
;     APPLICANT:  David Segev
;     TITLE OF INVENTION:  No. 5843650-enzymatic method for detecting nucleic acid seq
;     NUMBER OF SEQUENCES:  16
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Mark M. Friedman c/o Robert Sheinbein
;       STREET:  2940 Birchtree space lane
;       CITY:  Silver Spring
;       STATE:  Maryland
;       COUNTRY:  United States of America
;       ZIP:  20906
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  1.44 megabyte, 3.5" microdisk
;       COMPUTER:  Twinhead Slimnote-890TX
;       OPERATING SYSTEM:  MS DOS version 6.2,
;       OPERATING SYSTEM:  Windows version 3.11
;       SOFTWARE:  Word for Windows version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/431,527A
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Friedmam, Mark M.
;       REGISTRATION NUMBER:  33,883
;       REFERENCE/DOCKET NUMBER:  128/8
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  972-3-6938541
;       TELEFAX:  972-3-6938542
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  58
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-431-527A-7


  Query Match             2.1%;  Score 20.8;  DB 2;  Length 58;
  Best Local Similarity   60.7%;  Pred. No. 4.3e+03;
  Matches   34;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy     789 tggtgatgagaatggttctttaatttcagctatcaatgctgtaaaagatacaactg 844
           || || |||| |||| ||    ||   |||  | |||| | ||  | ||||  ||
Db      58 TGATGGTGAGGATGGGCCTCCGGTTCATGCTGCCCATGCAGGAACTGTTACACATG 3


RESULT  11
US-09-214-278-30/c
; Sequence 30, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
;  APPLICANT:  Sakano, Seiji
;  APPLICANT:  Itoh, Akira
```

```
;  TITLE OF INVENTION:  DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
;  FILE REFERENCE:  KP-8576
;  CURRENT APPLICATION NUMBER: US/09/214,278
;  CURRENT FILING DATE:  1999-01-26
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 30
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic DNA
US-09-214-278-30


  Query Match             2.1%;  Score 20.8;  DB 4;  Length 58;
  Best Local Similarity   78.1%;  Pred. No. 4.3e+03;
  Matches   25;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy    769 ggtaaggttgattattcagatggtgatgagaa 800
          || ||||  ||||||  ||||| ||||||  ||
Db     42 GGCAAGGAGGATTATAAAGATGATGATGATAA 11


RESULT  12
US-08-886-967-3/c
; Sequence 3, Application US/08886967
; Patent No. 6068993
;  GENERAL INFORMATION:
;    APPLICANT:  ASTOLFI, SPARTACO
;    APPLICANT:  DE LIMA, BEATRIZ D.
;    APPLICANT:  THIEMANN, JOSEF E.
;    APPLICANT:  TUNES DE SOUSA, HELOISA R.
;    APPLICANT:  VILELA, LUCIANO
;    TITLE OF INVENTION:  VECTOR FOR EXPRESSION OF HETEROLOGOUS
;    TITLE OF INVENTION:  PROTEIN AND METHODS FOR EXTRACTING RECOMBINANT PROTEIN AND
;    TITLE OF INVENTION:  FOR PURIFYING ISOLATED RECOMBINANT INSULIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FROMMER LAWRENCE & HAUG LLP
;      STREET:  745 FIFTH AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10151
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/886,967
;      FILING DATE:  02-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HAUG, EDGAR H.
;      REGISTRATION NUMBER:  29,309
;      REFERENCE/DOCKET NUMBER:  540519-2003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-588-0800
;      TELEFAX:  212-588-0500
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-886-967-3


  Query Match             2.1%;  Score 20.6;  DB 3;  Length 52;
  Best Local Similarity   67.4%;  Pred. No. 4.6e+03;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    400 ttaagtggtggttttaccaatcaagaattccaaatcggttcaa 442
          |||||||||   | ||   |  ||||||| ||| | | || ||
Db     52 TTAAGTGGTACCTATATAGAGGAAGAATTTCAATTTGTTTTAA 10


RESULT  13
US-09-306-949-3/c
; Sequence 3, Application US/09306949
; Patent No. 6281329
;  GENERAL INFORMATION:
;    APPLICANT:  ASTOLFI, SPARTACO
;    APPLICANT:  DE LIMA, BEATRIZ D.
;    APPLICANT:  THIEMANN, JOSEF E.
;    APPLICANT:  TUNES DE SOUSA, HELOISA R.
;    APPLICANT:  VILELA, LUCIANO
;    TITLE OF INVENTION:  VECTOR FOR EXPRESSION OF HETEROLOGOUS
;    TITLE OF INVENTION:  PROTEIN AND METHODS FOR EXTRACTING RECOMBINANT PROTEIN AND
;    TITLE OF INVENTION:  FOR PURIFYING ISOLATED RECOMBINANT INSULIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FROMMER LAWRENCE & HAUG LLP
;      STREET:  745 FIFTH AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10151
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/306,949
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/886,967
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HAUG, EDGAR H.
;      REGISTRATION NUMBER:  29,309
;      REFERENCE/DOCKET NUMBER:  540519-2003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-588-0800
;      TELEFAX:  212-588-0500
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  52 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-306-949-3


  Query Match             2.1%;  Score 20.6;  DB 4;  Length 52;
  Best Local Similarity   67.4%;  Pred. No. 4.6e+03;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    400 ttaagtggtggttttaccaatcaagaattccaaatcggttcaa 442
          |||||||||   | ||   |  ||||||| ||| | | || ||
Db     52 TTAAGTGGTACCTATATAGAGGAAGAATTTCAATTTGTTTTAA 10


RESULT  14
US-08-874-825-118/c
; Sequence 118, Application US/08874825
; Patent No. 6057101
;  GENERAL INFORMATION:
;    APPLICANT:  Nandabalan, Krishnan
;    APPLICANT:  Rothberg, Jonathan
```

```
;    APPLICANT:  Yang, Meijia
;    APPLICANT:  Knight, James
;    APPLICANT:  Kalbfleisch, Theodore
;    TITLE OF INVENTION:  IDENTIFICATION AND COMPARISON OF
;    TITLE OF INVENTION:  PROTEIN-PROTEIN INTERACTIONS THAT OCCUR IN POPULATIONS
;    TITLE OF INVENTION:  AND IDENTIFICATION OF INHIBITORS OF THESE INTERACTORS
;    NUMBER OF SEQUENCES:  122
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036/2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/874,825
;      FILING DATE:  13-JUN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/663,824
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  7934-045
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  118:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-874-825-118


  Query Match             2.0%;  Score 20;  DB 3;  Length 39;
  Best Local Similarity   82.1%;  Pred. No. 6e+03;
  Matches   23;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    768 aggtaaggttgattattcagatggtgat 795
          |||| |||||| ||  ||||||||| ||
Db     39 AGGTCAGGTTGCTTTCTCAGATGGTCAT 12


RESULT  15
US-08-450-905B-7/c
; Sequence 7, Application US/08450905B
; Patent No. 5856301
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Stem Cell Inhibiting Proteins
;    NUMBER OF SEQUENCES:  178
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HALE and DORR
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  MA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,905B
;      FILING DATE:  26-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/982,759
;      FILING DATE:  08-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9127319.3
;      FILING DATE:  23-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9221587.0
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BAKER, HOLLIE L.
;      REGISTRATION NUMBER:  31,321
;      REFERENCE/DOCKET NUMBER:  102.378.120DV-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-526-6110
;      TELEFAX:  617-526-5000
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..50
;      OTHER INFORMATION:  /product= "OLIGOMER FOR
;      OTHER INFORMATION:  CONSTRUCTION OF SYNTHETIC LD78 GENE"
US-08-450-905B-7


  Query Match             2.0%;  Score 20;  DB 2;  Length 50;
  Best Local Similarity   72.2%;  Pred. No. 6.6e+03;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    368 caaataccacttcatttaatggcaaacaacttttaa 403
          ||||| ||||   |||| || ||  || ||||| ||
Db     36 CAAATTCCACAAAATTTCATTGCTGACTACTTTGAA 1


Search completed: April 17, 2002, 02:18:56
Job time: 3297 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 17, 2002, 01:16:14 ; Search time 1464.96 Seconds
(without alignments)
7327.868 Million cell updates/sec

Title: US-09-439-311-1
Perfect score: 999
Sequence: 1 attaacacaaatgttgcagc.....ttaaaaatgatggtagagat 999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0
Maximum DB seq length: 60

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	23.6	2.4	50	10	AU104183	AU104183 AU104183
	2	21.6	2.2	60	10	AA608932	AA608932 af03d01.s
	3	21	2.1	52	10	AA122474	AA122474 mq27h10.r
c	4	20.8	2.1	59	10	AW687832	AW687832 NF014A01R
c	5	20.8	2.1	60	11	BF638864	BF638864 NF060H02P
c	6	20.6	2.1	58	10	AA874675	AA874675 vw84c07.r
c	7	20.4	2.0	51	13	AZ812517	AZ812517 2M0079C15
	8	20.4	2.0	55	10	AU014315	AU014315 AU014315
c	9	20.2	2.0	50	10	AU103731	AU103731 AU103731
	10	20.2	2.0	54	13	AZ954544	AZ954544 2M0220G15
c	11	20.2	2.0	58	11	BF131272	BF131272 601819529
c	12	20.2	2.0	59	11	W38842	W38842 zb28c08.r1
	13	20	2.0	54	13	AZ434413	AZ434413 1M0220I18
	14	20	2.0	56	10	AW780772	AW780772 sl85c06.y
	15	20	2.0	56	13	AF149647	AF149647 AF149647
	16	20	2.0	56	13	AZ451302	AZ451302 1M0250D12
c	17	20	2.0	58	10	AA164130	AA164130 mq84f08.r
	18	20	2.0	60	10	AA878830	AA878830 of83f08.s
c	19	20	2.0	60	13	AZ778864	AZ778864 2M0014P24
	20	20	2.0	60	13	B44939	B44939 HS-1060-A2-
	21	19.8	2.0	50	11	BG939021	BG939021 cn30a04.y
	22	19.8	2.0	58	11	BG058837	BG058837 nag43c06.
	23	19.8	2.0	60	10	AU060311	AU060311 AU060311
c	24	19.6	2.0	50	13	AZ397298	AZ397298 1M0162A13
	25	19.6	2.0	54	10	AW307272	AW307272 sf54h07.y
	26	19.6	2.0	59	13	AZ626633	AZ626633 1M0467C04
c	27	19.4	1.9	50	10	AU107058	AU107058 AU107058
c	28	19.4	1.9	52	10	AW687820	AW687820 NF013G12R
c	29	19.4	1.9	53	11	BG370297	BG370297 nai29h10.
c	30	19.4	1.9	56	10	AW513614	AW513614 xo47d06.x
	31	19.4	1.9	56	13	AZ608724	AZ608724 1M0433C09
c	32	19.4	1.9	58	10	AA190288	AA190288 mt93g05.r
c	33	19.4	1.9	58	11	T18527	T18527 hbc2089 Hum
c	34	19.4	1.9	58	11	T18565	T18565 hbc2087 Hum
c	35	19.4	1.9	59	10	AI855540	AI855540 sc20e08.y
	36	19.2	1.9	51	11	D18206	D18206 MUSGS00476
	37	19.2	1.9	52	10	AA492773	AA492773 vi76h09.r
	38	19.2	1.9	54	13	AZ783521	AZ783521 2M0025P17
c	39	19	1.9	50	11	BG271613	BG271613 nai59h03.
c	40	19	1.9	52	10	AW396361	AW396361 sh27b06.y
	41	19	1.9	53	11	BG408927	BG408927 gb83c07.y
c	42	19	1.9	54	13	AZ345736	AZ345736 1M0080M03
	43	19	1.9	60	11	C01215	C01215 HUMGS000792
	44	19	1.9	60	11	BF638586	BF638586 NF054G03P
	45	18.8	1.9	50	10	AU104948	AU104948 AU104948

ALIGNMENTS

RESULT 1
AU104183/c
LOCUS AU104183 50 bp mRNA EST 05-APR-2001
DEFINITION AU104183 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone KAT06093, mRNA sequence.
ACCESSION AU104183
VERSION AU104183.1 GI:13553704
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo,K., Suyama,A. and Sugano,S.
TITLE Fine Structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL Unpublished (2001)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT06093"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 12 a 12 c 4 g 22 t
ORIGIN

```
  Query Match             2.4%;  Score 23.6;  DB 10;  Length 50;
  Best Local Similarity   69.6%;  Pred. No. 1.5e+05;
  Matches   32;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    902 gagggattaaaatcacaggtagcataggtgtaggagctggtatatt 947
          ||||| ||||||| | || ||| |  ||  ||  |||| | || ||
Db     46 GAGGGTTTAAAATAAAAGTTAGAACTGGAATAAAAGCTAGCATGTT 1


RESULT   2
AA608932
LOCUS       AA608932       60 bp    mRNA            EST       02-MAR-1998
DEFINITION  af03d01.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1030561
            3' similar to gb:M10058 ASIALOGLYCOPROTEIN RECEPTOR 1 (HUMAN);;
            mRNA sequence.
ACCESSION   AA608932
VERSION     AA608932.1  GI:2457360
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 1146   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1030561"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       16 a     20 c     11 g     13 t
ORIGIN


  Query Match             2.2%;  Score 21.6;  DB 10;  Length 60;
  Best Local Similarity   63.5%;  Pred. No. 4.5e+05;
  Matches   33;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    205 ggtatcttgcaaactgcagataaggctatggatgagcaacttaaaatcttag 256
          |||| | ||| |||||||||| || || | || |    | ||    | ||| |
Db      9 GGTAACCTGCCAACTGCAGAAAACGCCACGGGTTTCAAGCTCCTCACCTTCG 60


RESULT   3
AA122474
LOCUS       AA122474       52 bp    mRNA            EST       13-FEB-1997
DEFINITION  mq27h10.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:580003 5'
            similar to SW:COX3_MOUSE P00416 CYTOCHROME C OXIDASE POLYPEPTIDE
            III ;, mRNA sequence.
ACCESSION   AA122474
VERSION     AA122474.1  GI:1681537
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:354651
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:580003"
                     /clone_lib="Barstead MPLRB1"
                     /sex="mixed"
                     /tissue_type="Kidney"
                     /dev_stage="6 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [CATGGATTCGGTACC], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT       20 a      7 c      8 g     17 t
ORIGIN


  Query Match             2.1%;  Score 21;  DB 10;  Length 52;
  Best Local Similarity   73.0%;  Pred. No. 6.2e+05;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    810 aatttcagctatcaatgctgtaaaagatacaactggt 846
          ||||||| | || |||||| || ||| | |||  | |
Db      4 AATTTCAACAATTAATGCTCTAGAAGTTGCAATAGTT 40


RESULT   4
AW687832/c
LOCUS       AW687832       59 bp    mRNA            EST       20-DEC-2000
DEFINITION  NF014A01RT1F1003 Developing root Medicago truncatula cDNA clone
            NF014A01RT 5', mRNA sequence.
```

```
ACCESSION   AW687832
VERSION     AW687832.2  GI:11929203
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
            Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May
            ,G.D. and Paiva,N.L.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula root library
  JOURNAL   Unpublished (2000)
COMMENT     On Apr 14, 2000 this sequence version replaced gi:7562568.
            Contact: Paiva NL
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7317
            Fax: 580 221 7380
            Email: nlpaiva@noble.org
            Insert Length: 721   Std Error: 0.00
            Plate: 014  row: A  column: 01
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF014A01RT"
                     /clone_lib="Developing root"
                     /tissue_type="root"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Total RNA was extracted from
                     non-nodulated roots of plants grown in 1 mM nitrate
                     medium.  Samples were taken at four time points
                     (approximately two days, one, two and six weeks post
                     germination) representing early seedling growth to
                     nitrogen limitation."
BASE COUNT       19 a      9 c     13 g     18 t
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 10;  Length 59;
  Best Local Similarity   70.0%;  Pred. No. 6.9e+05;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    822 caatgctgtaaaagatacaactggtgttcaagcctctaaa 861
          || |||  ||  ||| ||||| ||||||||  || | ||
Db     54 CATTGCAATAGCAGAAACAACAGGTGTTCATAACTTTCAA 15


RESULT   5
BF638864/c
LOCUS       BF638864       60 bp    mRNA            EST       19-DEC-2000
DEFINITION  NF060H02PL1F1026 Phosphate starved leaf Medicago truncatula cDNA
            clone NF060H02PL 5', mRNA sequence.
ACCESSION   BF638864
VERSION     BF638864.1  GI:11903022
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
            ,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula phosphate-starved leaf library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Harrison MJ
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7325
            Fax: 580 221 7380
            Email: mjharrison@noble.org
            Insert Length: 60   Std Error: 0.00
            Plate: 060  row: H  column: 02
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF060H02PL"
                     /clone_lib="Phosphate starved leaf"
                     /tissue_type="leaf"
                     /dev_stage="trifoliate"
                     /note="Vector: Lambda Zap; At the trifoliate stage, M.
                     truncatula plants were transplanted to phosphate-free sand
                     and grown for a further 30 days.  During this 30 day
                     period, the plants were fertilized twice weekly with 1/2
                     Hoaglands solution containing only 20uM potassium
                     phosphate.  RNA was prepared from above ground tissues."
BASE COUNT       15 a     17 c     12 g      9 t      7 others
ORIGIN


  Query Match             2.1%;  Score 20.8;  DB 11;  Length 60;
  Best Local Similarity   59.6%;  Pred. No. 6.9e+05;
  Matches   28;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     76 tcaagactcagttcaggtcttagaatcaactccgcagcagatgatgc 122
          ||| ||||  ||| |||| | |    || ||   | |  | || |||
Db     47 TCAGGACTNNGTTNAGGTTTGATGGNCAGCTANACNGNTGCTGGTGC 1


RESULT   6
AA874675/c
LOCUS       AA874675       58 bp    mRNA            EST       19-MAR-1998
DEFINITION  vw84c07.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1261644 5' similar to TR:O03713 O03713 CYTOCHOROME B ;, mRNA
            sequence.
ACCESSION   AA874675
VERSION     AA874675.1  GI:2978571
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:664196
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
```

```
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1261644"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     18 c      2 g     19 t
ORIGIN


  Query Match             2.1%;  Score 20.6;  DB 10;  Length 58;
  Best Local Similarity   62.7%;  Pred. No. 7.7e+05;
  Matches   32;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    208 atcttgcaaactgcagataaggctatggatgagcaacttaaaatcttagat 258
          || ||| |    |||| || | | ||| |||||    | | ||| ||| ||
Db     51 ATGTTGGATGGGGCAGGTAGGTCAATGAATGAGTGGTTAATAATTTTAAAT 1


RESULT   7
AZ812517/c
LOCUS       AZ812517       51 bp    DNA             GSS       20-FEB-2001
DEFINITION  2M0079C15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0079C15 F, DNA sequence.
ACCESSION   AZ812517
VERSION     AZ812517.1  GI:12981841
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0079  row: C  column: 15
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 51.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0079C15"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        9 a     18 c      11 g     13 t
ORIGIN


  Query Match             2.0%;  Score 20.4;  DB 13;  Length 51;
  Best Local Similarity   71.1%;  Pred. No. 8.6e+05;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    127 gggatggcgatagcagatagtttaagatctcaggcaaa 164
          ||||   |||  ||||| |||||  || ||||||  ||
Db     50 GGGAGCCCGAAGGCAGAAAGTTTTTGAGCTCAGGGCAA 13


RESULT   8
AU014315
LOCUS       AU014315       55 bp    mRNA            EST       03-AUG-1998
DEFINITION  AU014315 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc09537, mRNA sequence.
ACCESSION   AU014315
VERSION     AU014315.1  GI:3369106
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc09537"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
BASE COUNT       24 a      2 c       9 g     20 t
ORIGIN


  Query Match             2.0%;  Score 20.4;  DB 10;  Length 55;
```

```
  Best Local Similarity   65.2%;  Pred. No. 8.6e+05;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    172 ggtcaggctatatctaatggtaatgatgctttaggtatcttgcaaa 217
          ||| | |  |||  ||  | |||| | | |||| ||||  ||| ||
Db      2 GGTTATGTAATAGATATAGTTAATAACGATTTATGTATTGTGCTAA 47


RESULT  9
AU103731/c
LOCUS       AU103731       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU103731 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            COL04386, mRNA sequence.
ACCESSION   AU103731
VERSION     AU103731.1  GI:13553252
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COL04386"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        6 a     12 c     13 g     19 t
ORIGIN


  Query Match             2.0%;  Score 20.2;  DB 10;  Length 50;
  Best Local Similarity   63.3%;  Pred. No. 9.6e+05;
  Matches   31;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    636 tttggctgaagagatcaacagaaatgcagataaaacaggaattcgtgca 684
          ||||||  | || | | ||  | ||| | | || |||  ||| |  |||
Db     50 TTTGGCGCAGGAAAGCGACCAACATGGACAGAAGACAACAATCCACGCA 2


RESULT  10
AZ954544
LOCUS       AZ954544       54 bp    DNA             GSS       27-APR-2001
DEFINITION  2M0220G15F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0220G15 F, DNA sequence.
ACCESSION   AZ954544
VERSION     AZ954544.1  GI:13825771
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0220  row: G  column: 15
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 54.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0220G15"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       18 a      9 c     12 g     15 t
ORIGIN


  Query Match             2.0%;  Score 20.2;  DB 13;  Length 54;
  Best Local Similarity   63.3%;  Pred. No. 9.6e+05;
  Matches   31;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    846 tgttcaagcctctaaagatgaaaatggtaaacttgttcttacttcggcc 894
          |||  |||  | ||||||||||| |     ||| || ||||  ||   ||
Db      6 TGTCTAAGTGTGTAAAGATGAAGAATAGGAACCTGCTCTTCTTTACACC 54


RESULT  11
BF131272/c
LOCUS       BF131272       58 bp    mRNA            EST       24-OCT-2000
DEFINITION  601819529F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4051302 5',
            mRNA sequence.
ACCESSION   BF131272
VERSION     BF131272.1  GI:10970312
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
            Plate: LLCM886  row: j  column: 07.
FEATURES              Location/Qualifiers
     source           1. .58
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4051302"
                      /clone_lib="NIH_MGC_58"
                      /tissue_type="hypernephroma"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                      SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                      Double-stranded cDNA was prepared from cell line RNA.  5'
                      and 3' adaptors were used in cloning as follows: 5'
                      adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                      sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                      (where B = A, C, or G and N = A, C,  G, or T).   Average
                      insert size 1.35 kb (range 0.9-4.0 kb).  15/15 colonies
                      contained  inserts by PCR. This library was enriched for
                      full-length clones and was constructed by Clontech
                      Laboratories (Palo Alto, CA)."
BASE COUNT       22 a      9 c      0 g      26 t      1 others
ORIGIN


  Query Match              2.0%;  Score 20.2;  DB 11;  Length 58;
  Best Local Similarity   63.3%;  Pred. No. 9.6e+05;
  Matches   31;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     543 tgtaggacttactattaaaaactacaacggtatcgaagattttaaattt 591
           ||||  |   |  | ||| || || || |||||  |||  | |||||
Db      49 TGTATAATAGATAAGTAATAATTAAAATGGTATGTAAGGATGTAAATAT 1


RESULT  12
W38842/c
LOCUS       W38842         59 bp    mRNA            EST       15-MAY-1996
DEFINITION  zb28c08.r1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:304910 5' similar to gb:M90516
            GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINOTRANSFERASE (HUMAN);, mRNA
            sequence.
ACCESSION   W38842
VERSION     W38842.1  GI:1320547
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: mob.REGA+ET
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .59
                      /organism="Homo sapiens"
                      /db_xref="GDB:1248320"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:304910"
                      /clone_lib="Soares_parathyroid_tumor_NbHPA"
                      /tissue_type="parathyroid tumor"
                      /dev_stage="adult"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                      ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                      RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                      primer
                      [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTT
                      TTTTT-3'], double-stranded cDNA was size selected, ligated
                      to Eco RI adapters (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of a modified pT7T3
                      vector (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                      adenomas was kindly provided by Dr. Stephen Marx, National
                      Institute of Diabetes and Digestive and Kidney Diseases,
                      NIH."
BASE COUNT       23 a      5 c      10 g      19 t      2 others
ORIGIN


  Query Match              2.0%;  Score 20.2;  DB 11;  Length 59;
  Best Local Similarity   63.3%;  Pred. No. 9.6e+05;
  Matches   31;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     550 cttactattaaaaactacaacggtatcgaagattttaaatttgatagtg 598
           ||| .| || ||||  |||||   |  |  | | |||||||  ||| ||
Db      51 CTTGTTCTTTAAAATCACAACTACAGAGCTGCTATTAAATTCCATATTG 3


RESULT  13
AZ434413
LOCUS       AZ434413       54 bp    DNA             GSS       03-OCT-2000
DEFINITION  1M0220I18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0220I18 R, DNA sequence.
ACCESSION   AZ434413
VERSION     AZ434413.1  GI:10558426
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0220  row: I  column: 18
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 54.
FEATURES              Location/Qualifiers
     source           1. .54
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
```

/db_xref="taxon:10090"
/clone="UUGC1M0220I18"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 19 a 11 c 6 g 18 t
ORIGIN

Query Match 2.0%; Score 20; DB 13; Length 54;
Best Local Similarity 61.5%; Pred. No. 1.1e+06;
Matches 32; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 836 atacaactggtgttcaagcctctaaagatgaaaatggtaaacttgttcttac 887
|| ||| | || | | ||||| | ||| ||||| | || | || ||
Db 2 ATTCAAAATGCTTTTATGTCTCTAGAAATGTAAATGATCAAATGTCTCCAAC 53

RESULT 14
AW780772
LOCUS AW780772 56 bp mRNA EST 12-MAY-2000
DEFINITION s185c06.y1 Gm-c1037 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1037-803 5', mRNA sequence.
ACCESSION AW780772
VERSION AW780772.1 GI:7795447
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 56)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Putative full length read
vector to vector length is 57.
FEATURES Location/Qualifiers
source 1. .56
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1037-803"
/clone_lib="Gm-c1037"
/tissue_type="fully expanded leaves of greenhouse grown plants"
/dev_stage="2 week old"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This cDNA library was constructed from mRNA isolated from fully expanded leaves of greenhouse grown plants that were 2 weeks old. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restrictions site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electro- Max DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana- Champaign. email: l-vodkin@uiuc.edu"
BASE COUNT 15 a 2 c 11 g 28 t
ORIGIN

Query Match 2.0%; Score 20; DB 10; Length 56;
Best Local Similarity 65.9%; Pred. No. 1.1e+06;
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 846 tgttcaagcctctaaagatgaaaatggtaaacttgttcttactt 889
| || |||| ||| ||||| || |||| || | ||| ||
Db 3 TTTTGAAGCTGCTATAGATGTTAAATGTAATTTTAGTTTTAGTT 46

RESULT 15
AF149647
LOCUS AF149647 56 bp DNA GSS 12-JUN-2000
DEFINITION AF149647 Human chromosome 18q21 from exon-trapping Homo sapiens genomic clone 5m10, DNA sequence.
ACCESSION AF149647
VERSION AF149647.1 GI:8485985
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 56)
AUTHORS Chen,H., Huo,Y., Patel,S., Zhu,X., Swift-Scanlan,T., Reeves,R.H., DePaulo,R. Jr., Ross,C.A. and McInnis,M.G.
TITLE Gene identification using exon amplification on human chromosome 18q21: implications for bipolar disorder
JOURNAL Mol. Psychiatry 5 (5), 502-509 (2000)
MEDLINE 20485132
COMMENT Contact: Chen H
Psychiatry and Behavioral Sciences
Johns Hopkins University School of Medicine
600 N. Wolfe Street, Baltimore, MD 21287, USA
Email: hc@welchlink.welch.jhu.edu
Class: exon-trapped.
FEATURES Location/Qualifiers
source 1. .56
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="18q21"
/clone="5m10"
/clone_lib="Human chromosome 18q21 from exon-trapping"
BASE COUNT 24 a 12 c 7 g 13 t
ORIGIN

Query Match 2.0%; Score 20; DB 13; Length 56;
Best Local Similarity 61.5%; Pred. No. 1.1e+06;
Matches 32; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy     335 tcaaccgtttgatggaagaacttgataatatcgcaaataccacttcatttaa 386
           ||| | |    | | ||  | |||  |||| | ||||| |  | |||||| |
Db       4 TCAGCAGCCAAAAGAAAAGATTTGTCAATACCACAAATGCATCATCATTTTA 55
```

Search completed: April 17, 2002, 02:14:13
Job time: 3479 sec

OM of: US-09-439-311-2 to: N_Geneseq_1101:* out_format : pfs

Date: Apr 17, 2002 3:13 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09439311/runat_16042002_134011_11729/app_query.fasta_1.395
-DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=60 -USER=US09439311_@CGN1_1_215
-NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-439-311-2
Query length: 333
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 170.970000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation ..
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA86891	-	50.00	100.69	1.6e+03	30	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA86892	+	50.00	100.69	1.6e+03	30	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA93188	-	50.00	100.69	1.6e+03	30	! C
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA93190	+	50.00	100.69	1.6e+03	30	! C
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV23196	+	46.00	88.52	7.4e+03	59	! L
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX34813	+	43.00	85.32	1.1e+04	51	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX19519	+	43.00	85.32	1.1e+04	51	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV00234	+	41.00	82.48	1.6e+04	50	! T
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ96924	+	40.00	79.47	2.4e+04	59	! S
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAT00254	+	40.00	79.32	2.4e+04	60	! T
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF70806	+	40.00	79.32	2.4e+04	60	! T
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI30690	+	39.00	83.80	1.4e+04	31	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV56429	-	39.00	80.03	2.2e+04	47	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX21884	-	39.00	80.03	2.2e+04	47	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV69197	-	39.00	80.03	2.2e+04	47	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA97171	-	39.00	80.03	2.2e+04	47	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA08317	-	39.00	80.03	2.2e+04	47	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ24343	-	39.00	80.03	2.2e+04	47	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH36472	-	39.00	77.97	2.9e+04	59	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA27149	-	38.00	81.73	1.8e+04	33	! C
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC62765	-	38.00	79.98	2.2e+04	40	! S
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH79904	+	38.00	77.78	2.9e+04	51	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAQ95069	-	38.00	77.43	3.1e+04	53	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX05302	-	38.00	77.43	3.1e+04	53	! C
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAT23184	-	38.00	76.46	3.5e+04	59	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ30956	+	38.00	76.31	3.5e+04	60	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT60508	-	37.00	83.11	1.5e+04	24	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV31446	-	37.00	83.11	1.5e+04	24	! C
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV25942	-	37.00	83.11	1.5e+04	24	! C
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV20847	-	37.00	83.11	1.5e+04	24	! C
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT63391	-	37.00	79.43	2.4e+04	36	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ96013	+	37.00	78.48	2.7e+04	40	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA73050	+	37.00	76.63	3.4e+04	49	! T
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA77507	-	37.00	76.27	3.5e+04	51	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA71013	-	37.00	75.42	4.0e+04	56	! M
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA71017	-	37.00	75.42	4.0e+04	56	! M
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX57663	+	37.00	75.11	4.1e+04	58	! M
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV18769	+	36.50	74.51	4.4e+04	57	! P
/SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ56490	-	36.00	74.76	4.3e+04	51	! D
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT73784	-	36.00	74.76	4.3e+04	51	! M
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH79670	+	36.00	74.76	4.3e+04	51	! H
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH79776	-	36.00	74.76	4.3e+04	51	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD06898	-	36.00	74.76	4.3e+04	51	
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH38320	-	36.00	74.76	4.3e+04	51	
/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAQ97415	+	36.00	74.41	4.5e+04	53	

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA86891

seq_documentation_block:
ID AAA86891 standard; DNA; 30 BP.
XX
AC AAA86891;
XX
DT 15-JAN-2001 (first entry)
XX
DE Probe to Campylobacter jejuni.
XX
KW Detection; nucleic acid hybrid; depolymerisation; analysis; SNP;
KW single nucleotide polymorphism; identification; viral load; probe;
KW genotyping; medical marker diagnostic; primer; target; mutation;
KW genetic disease; ss.
XX
OS Campylobacter jejuni.
XX
PN WO200049180-A1.
XX
PD 24-AUG-2000.
XX
PF 18-FEB-2000; 2000WO-US04242.
XX
PR 18-FEB-1999; 99US-0252436.
PR 21-JUL-1999; 99US-0358972.
PR 25-AUG-1999; 99US-0383316.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Shultz JW, Lewis MK, Leippe D, Mandrekar M, Kephart D, Rhodes RB;
PI Andrews CA, Hartnett JR, Gu T, Olson RJ, Wood KV, Welch R;
XX
DR WPI; 2000-565377/52.
XX
PT Determining presence or absence of a predetermined endogenous nucleic
PT acid sequence by using an enzyme that depolymerizes the 3' end of an
PT oligonucleotide probe hybridized to a target sequence to release
PT identifier nucleotides -
XX
PS Example; Page 321; 389pp; English.
XX
CC The present invention describes a method (M1) for determining the
CC presence or absence of a predetermined endogenous nucleic acid target
CC sequence (ENAT). The method comprises hybridising a probe having an
CC identifier nucleotide (IN) with ENAT which is treated with an enzyme
CC that depolymerises the 3' end of hybridised NA to release the INs.
CC M1 is used for determining the number of known sequence repeats present
CC in a nucleic acid target sequence in a nucleic acid sample. The method
CC is also useful for determining whether a nucleic acid target sequence in
CC a sample is an allele from a homozygous or heterozygous locus. The
CC method is also useful for detection of mutations, translocations and
CC SNPs in nucleic acids (including those associated with genetic disease),
CC determination of viral load, species identification, sample
CC contamination, and analysis of forensic samples. AAA86791 to AAA87079
CC and AAB12817 represent sequence which are used in the exemplification of
CC the present invention.
CC N.B. There is a discrepancy between the SEQ ID NO: and sequences given
CC in the examples, and the SEQ ID NO: and sequences given in the sequence
CC listing from the present invention.
XX
SQ Sequence 30 BP; 5 A; 5 C; 4 G; 16 T; 0 other;

alignment_scores:
Quality: 50.00 Length: 10
Ratio: 5.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```
alignment_block:
 US-09-439-311-2 x AAA86891/rev   ..

 Align seg 1/1  to reverse of: AAA86891  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
      30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA86892

seq_documentation_block:
ID   AAA86892 standard; DNA; 30 BP.
XX
AC   AAA86892;
XX
DT   15-JAN-2001  (first entry)
XX
DE   Probe to Campylobacter jejuni.
XX
KW   Detection; nucleic acid hybrid; depolymerisation; analysis; SNP;
KW   single nucleotide polymorphism; identification; viral load; probe;
KW   genotyping; medical marker diagnostic; primer; target; mutation;
KW   genetic disease; ss.
XX
OS   Campylobacter jejuni.
XX
PN   WO200049180-A1.
XX
PD   24-AUG-2000.
XX
PF   18-FEB-2000; 2000WO-US04242.
XX
PR   18-FEB-1999;   99US-0252436.
PR   21-JUL-1999;   99US-0358972.
PR   25-AUG-1999;   99US-0383316.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Shultz JW,  Lewis MK,  Leippe D,  Mandrekar M,  Kephart D,  Rhodes RB;
PI   Andrews CA,  Hartnett JR,  Gu T,  Olson RJ,  Wood KV,  Welch R;
XX
DR   WPI; 2000-565377/52.
XX
PT   Determining presence or absence of a predetermined endogenous nucleic
PT   acid sequence by using an enzyme that depolymerizes the 3' end of an
PT   oligonucleotide probe hybridized to a target sequence to release
PT   identifier nucleotides  -
XX
PS   Example; Page 321; 389pp; English.
XX
CC   The present invention describes a method (M1) for determining the
CC   presence or absence of a predetermined endogenous nucleic acid target
CC   sequence (ENAT). The method comprises hybridising a probe having an
CC   identifier nucleotide (IN) with ENAT which is treated with an enzyme
CC   that depolymerises the 3' end of hybridised NA to release the INs.
CC   M1 is used for determining the number of known sequence repeats present
CC   in a nucleic acid target sequence in a nucleic acid sample. The method
CC   is also useful for determining whether a nucleic acid target sequence in
CC   a sample is an allele from a homozygous or heterozygous locus. The
CC   method is also useful for detection of mutations, translocations and
CC   SNPs in nucleic acids (including those associated with genetic disease),
CC   determination of viral load, species identification, sample
CC   contamination, and analysis of forensic samples. AAA86791 to AAA87079
CC   and AAB12817 represent sequence which are used in the exemplification of
CC   the present invention.
CC   N.B. There is a discrepancy between the SEQ ID NO: and sequences given
CC   in the examples, and the SEQ ID NO: and sequences given in the sequence
CC   listing from the present invention.
XX
SQ   Sequence 30 BP; 16 A; 4 C; 5 G; 5 T; 0 other;
```

```
alignment_scores:
            Quality:   50.00            Length:     10
              Ratio:  5.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-439-311-2 x AAA86892   ..

 Align seg 1/1  to: AAA86892  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
       1 CAAGATGGACAAAGTTTAAAAACAAGAACT 30

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA93188

seq_documentation_block:
ID   AAA93188 standard; DNA; 30 BP.
XX
AC   AAA93188;
XX
DT   11-JAN-2001  (first entry)
XX
DE   Campylobacter jejuni interrogation probe 11451.
XX
KW   Campylobacter jejuni; nucleic acid detection; genomic typing;
KW   mutation detection; viral load determination; species identification;
KW   forensic analysis; probe; ss.
XX
OS   Campylobacter jejuni.
XX
PN   WO200049179-A1.
XX
PD   24-AUG-2000.
XX
PF   18-FEB-2000; 2000WO-US04176.
XX
PR   18-FEB-1999;   99US-0252436.
PR   21-JUL-1999;   99US-0358972.
PR   27-SEP-1999;   99US-0406147.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Shultz JW,  Lewis MK,  Leippe D,  Mandrekar M,  Kephart D,  Rhodes RB;
PI   Andrews CA,  Hartnett JR,  Gu T,  Olson RJ,  Wood KV,  Welch R;
XX
DR   WPI; 2000-549282/50.
XX
PT   Detecting the presence of predetermined exogenous nucleic acid target
PT   sequence useful for e.g. genotyping, comprises depolymerizing the 3'
PT   end of an oligonucleotide probe hybridized to a nucleic acid target
PT   sequence  -
XX
PS   Claim 47; Page 187; 230pp; English.
XX
CC   The present sequence is an interrogation probe which was used to detect a
CC   segment of the genome of Campylobacter jejuni. This was performed as part
CC   of a method for determining the presence of a known exogenous nucleic
CC   acid target sequence in a nucleic acid sample. The method comprises
CC   admixing a treated sample with a depolymerising enzyme which releases one
CC   or more nucleotides from the 3'-end of a hybridised nucleic acid probe.
CC   The method is used for assaying nucleic acids for a particular native or
CC   mutant sequence, and for genomic typing. It is useful for detecting
CC   mutations, translocations, and single nucleotide polymorphisms,
CC   determination of viral load, species identification, detection of sample
CC   contamination, and analysis of forensic samples. Compared with previous
CC   methods of detecting nucleic acid hybrids, the new method has higher
CC   sensitivity without the need for radiochemicals or electrophoresis. It is
CC   quantitative, highly reproducible and can be automated. The method can
CC   reliably detect as few as 10 copies of a virus in a sample, and is
CC   capable of providing multiple analyses in a single assay (multiplex
CC   assay).
```

```
XX
SQ   Sequence 30 BP; 5 A; 5 C; 4 G; 16 T; 0 other;


alignment_scores:
            Quality:   50.00             Length:     10
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-439-311-2 x AAA93188/rev   ..

 Align seg 1/1  to reverse of: AAA93188  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
      30 CAAGATGGACAAAGTTTAAAAACAAGAACT 1

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA93190

seq_documentation_block:
ID   AAA93190 standard; DNA; 30 BP.
XX
AC   AAA93190;
XX
DT   11-JAN-2001  (first entry)
XX
DE   Campylobacter jejuni interrogation probe 11450.
XX
KW   Campylobacter jejuni; nucleic acid detection; genomic typing;
KW   mutation detection; viral load determination; species identification;
KW   forensic analysis; probe; ss.
XX
OS   Campylobacter jejuni.
XX
PN   WO200049179-A1.
XX
PD   24-AUG-2000.
XX
PF   18-FEB-2000; 2000WO-US04176.
XX
PR   18-FEB-1999;   99US-0252436.
PR   21-JUL-1999;   99US-0358972.
PR   27-SEP-1999;   99US-0406147.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Shultz JW,  Lewis MK,  Leippe D,  Mandrekar M,  Kephart D,  Rhodes RB;
PI   Andrews CA,  Hartnett JR,  Gu T,  Olson RJ,  Wood KV,  Welch R;
XX
DR   WPI; 2000-549282/50.
XX
PT   Detecting the presence of predetermined exogenous nucleic acid target
PT   sequence useful for e.g. genotyping, comprises depolymerizing the 3'
PT   end of an oligonucleotide probe hybridized to a nucleic acid target
PT   sequence  -
XX
PS   Claim 47; Page 187; 230pp; English.
XX
CC   The present sequence is an interrogation probe which was used to detect a
CC   segment of the genome of Campylobacter jejuni. This was performed as part
CC   of a method for determining the presence of a known exogenous nucleic
CC   acid target sequence in a nucleic acid sample. The method comprises
CC   admixing a treated sample with a depolymerising enzyme which releases one
CC   or more nucleotides from the 3'-end of a hybridised nucleic acid probe.
CC   The method is used for assaying nucleic acids for a particular native or
CC   mutant sequence, and for genomic typing. It is useful for detecting
CC   mutations, translocations, and single nucleotide polymorphisms,
CC   determination of viral load, species identification, detection of sample
CC   contamination, and analysis of forensic samples. Compared with previous
CC   methods of detecting nucleic acid hybrids, the new method has higher
CC   sensitivity without the need for radiochemicals or electrophoresis. It is
CC   quantitative, highly reproducible and can be automated. The method can
CC   reliably detect as few as 10 copies of a virus in a sample, and is
CC   capable of providing multiple analyses in a single assay (multiplex
CC   assay).
XX
SQ   Sequence 30 BP; 16 A; 4 C; 5 G; 5 T; 0 other;


alignment_scores:
            Quality:   50.00             Length:     10
              Ratio:  5.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-439-311-2 x AAA93190   ..

 Align seg 1/1  to: AAA93190  from: 1  to: 30

      97 GlnAspGlyGlnSerLeuLysThrArgThr 106
         ||||||||||||||||||||||||||||||
       1 CAAGATGGACAAAGTTTAAAAACAAGAACT 30

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV23196

seq_documentation_block:
ID   AAV23196 standard; DNA; 59 BP.
XX
AC   AAV23196;
XX
DT   28-JUL-1998  (first entry)
XX
DE   Lactococcus lactis constitutional promoter Cp29.
XX
KW   Lactococcus lactis; constitutional promoter; optimise; spacer;
KW   artificial promoter library; gene expression; ds.
XX
OS   Synthetic.
OS   Lactococcus lactis.
XX
FH   Key             Location/Qualifiers
FT   promoter        4..59
FT                   /*tag= a
FT                   /standard_name= "Constitutional promoter"
XX
PN   WO9807846-A1.
XX
PD   26-FEB-1998.
XX
PF   25-AUG-1997;   97WO-DK00342.
XX
PR   23-AUG-1996;   96DK-0000886.
XX
PA   (JENS/) JENSEN P R.
XX
PI   Hammer K,  Jensen PR;
XX
DR   WPI; 1998-179062/16.
XX
PT   New artificial promoter libraries - containing consensus promoter
PT   sequences and variable spacers, used to generate promoters for
PT   optimising expression of genes
XX
PS   Claim 28; Page 52; 89pp; English.
XX
CC   This is a Lactococcus lactis constitutional promoter sequence used in the
CC   construction of an artificial promoter library of the invention. The
CC   artificial promoter library for a selected organism or group of organisms
CC   comprise a mixture of double-stranded DNA fragments, the sense strands of
CC   which comprise at least half of two consensus sequences of efficient
CC   promoters from the organism or group of organisms and surrounding or
CC   intermediate nucleotide sequences (spacers) of variable length in which
CC   at least 7 nucleotides are selected randomly, with the proviso that
CC   previously known promoter sequences and promoter sequences isolated from
CC   natural sources are not included. This promoter library can be used in a
```

```
CC   method of optimising the expression of a gene in a microorganism. The
CC   method comprises selecting a set of promoters covering a range of
CC   promoter activities in relatively small steps of activity change from
CC   such an an artificial promoter library and cloning the set of promoters
CC   into the organism placing in each clone the gene under the control of
CC   at least one promoter from the set and growing the selected clones and
CC   screening them to find the one showing optimised flux of product
CC   formation. Promoters covering wide ranges of activities, including very
CC   strong promoters can be generated which can be used for optimising
CC   expression of genes.
XX
SQ   Sequence 59 BP; 15 A; 9 C; 16 G; 19 T; 0 other;


alignment_scores:
            Quality:   46.00             Length:     13
              Ratio:  4.600               Gaps:      0
 Percent Similarity:  76.923   Percent Identity: 76.923

alignment_block:
 US-09-439-311-2 x AAV23196   ..

 Align seg 1/1  to: AAV23196  from: 1  to: 59

     291 GlyLysLeuValLeuThrSerAlaAspGlyArgGlyIle 303
         |||||||||   |||||||||       ||||||||||||
       5 GGTAAGTTATTCTTGACATCTCAGGGGGGACGTGGTATA 43

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX34813

seq_documentation_block:
ID   AAX34813 standard; DNA; 51 BP.
XX
AC   AAX34813;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Human ZSIG-11 DNA specific primer ZC13735.
XX
KW   Secretory protein; ZSIG-11; ligand polypeptide; testis; endoprotease;
KW   prohormone convertase; fertility; therapeutic; human; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9916870-A1.
XX
PD   08-APR-1999.
XX
PF   29-SEP-1998;   98WO-US20449.
XX
PR   19-MAY-1998;   98US-0085966.
PR   29-SEP-1997;   97US-0060327.
PR   29-SEP-1997;   97US-0939897.
PR   19-MAY-1998;   98US-0081310.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO;
XX
DR   WPI; 1999-263692/22.
XX
PT   Polynucleotide encoding a human secretory protein, ZSIG-11.
XX
PS   Example 9; Page 108; 113pp; English.
XX
CC   The invention relates to a human secretory protein, ZSIG-11. Host cells
CC   containing a vector comprising the ZSIG-11 nucleic acid are used for the
CC   recombinant expression of the protein. ZSIG-11 is a novel ligand
CC   polypeptide and specific antibodies can be used to detect its presence
CC   in a biological sample. Probes derived from ZSIG-11 nucleotide sequences
CC   can also be used in detection of ZSIG-11 RNA. ZSIG-11 is expressed at
CC   high levels in testis, and could be used to identify/study prohormone
CC   convertases or endoproteases that exhibit testis specificity.
CC   Antagonists, including antibodies, are useful for inhibiting or
CC   eliminating the function of ZSIG-11.  It is possible that ZSIG-11 and
CC   its antagonists will be useful as fertility inducing therapeutics.
CC   Sequences AAX34800-21 represent PCR primers for amplifying the ZSIG-11
CC   DNA.
XX
SQ   Sequence 51 BP; 17 A; 5 C; 19 G; 10 T; 0 other;


alignment_scores:
            Quality:   43.00             Length:     16
              Ratio:  3.583               Gaps:      0
 Percent Similarity:  75.000   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x AAX34813   ..

 Align seg 1/1  to: AAX34813  from: 1  to: 51

     253 GlyValValIleGlyLysValAspTyrSerAspGlyAspGluAsnGly 268
         ||||||   :::   |||   ||||||:::|||   |||:::::|||
       1 GGTGTAAGCTTGGACAAGAGATTACAAGGACGATGATGACAAGGGT 48

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX19519

seq_documentation_block:
ID   AAX19519 standard; DNA; 51 BP.
XX
AC   AAX19519;
XX
DT   07-JUN-1999  (first entry)
XX
DE   Human lipocalin homologue zlipol PCR primer ZC13,735.
XX
KW   Human; lipocalin; testis; mammary gland; breast tumour; zlipol;
KW   breast cancer; emphysema; skin disease; reproduction; anti-inflammatory;
KW   antimicrobial; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9907740-A2.
XX
PD   18-FEB-1999.
XX
PF   06-AUG-1998;   98WO-US16425.
XX
PR   06-AUG-1997;   97US-0054867.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Conklin DC;
XX
DR   WPI; 1999-167367/14.
XX
PT   New lipocalin homologue designated zlipol - whose expression is
PT   restricted to testis and mammary gland tissues, particularly breast
PT   tumour tissue, used to, e.g. predict tumour aggressiveness.
XX
PS   Example 5; Page 89; 94pp; English.
XX
CC   The present sequence represents a PCR primer for lipocalin homologue,
CC   zlipol. The lipocalin homologue, zlipol, is specifically expressed in
CC   testis and mammary gland, particularly breast tumour tissue. Based on
CC   this tissue distribution, zlipol may be used as a diagnostic for breast
CC   carcinomas and as a tool for predicting tumour aggressiveness. Agonists
CC   can be used for transportation of small hydrophobic molecules either in
CC   vivo or in vitro, and so are useful in specifically promoting the growth
CC   and/or development of testis-specific cell lineages in culture. Zlipol
CC   can be used to identify inhibitors. Zlipol proteins can also be used to
CC   prepare antibodies (which can be linked to toxins), and can serve as
CC   immunogens. Zlipol proteins can be used as a delivery and encapsulation
```

```
CC   system to transport and/or stabilise small lipophilic molecules, e.g. to
CC   protect from gut pH and digestive enzymes. They can also be used to bind
CC   small fatty acids in blood or tissues to modulate their biological
CC   function, e.g. to transport retinoids or steroids to receptors, in
CC   particular as therapy for breast cancer, emphysema and diseases of the
CC   skin. They may also play an important role in reproduction. Other uses
CC   include anti-inflammatory responses, and antimicrobial activities.
CC   Zlipol nucleic acid sequences may be used for gene therapy to increase
CC   or inhibit zlipol activity, to derive probes and primers, to derive
CC   antisense sequences, and to detect genetic abnormalities.
XX
SQ   Sequence 51 BP; 17 A; 5 C; 19 G; 10 T; 0 other;


alignment_scores:
            Quality:   43.00              Length:     16
              Ratio:  3.583                Gaps:      0
 Percent Similarity:  75.000   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x AAX19519   ..

 Align seg 1/1  to: AAX19519  from: 1  to: 51

     253 GlyValValIleGlyLysValAspTyrSerAspGlyAspGluAsnGly 268
         ||||||   :::   |||   ||||||:::|||   |||::::::|||
       1 GGTGTAAGCTTGGACAAGAGAGATTACAAGGACGATGATGACAAGGGT 48

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV00234

seq_documentation_block:
ID   AAV00234 standard; DNA; 50 BP.
XX
AC   AAV00234;
XX
DT   08-JUN-1998  (first entry)
XX
DE   Tick vasoactive amine binding protein FS-HBP1 reverse PCR primer.
XX
KW   Female-specific vasoactive amine binding protein 1; FS-HCP1;
KW   histamine; serotonin; assay; antihistamine; anti-inflammatory;
KW   insect bite; snake bite; scorpion bite; dermatitis; vaccine;
KW   transgenic animal; tick; PCR; primer; ss.
XX
OS   Synthetic.
OS   Rhipicephalus appendiculatus.
XX
PN   WO9744451-A2.
XX
PD   27-NOV-1997.
XX
PF   19-MAY-1997;   97WO-GB01372.
XX
PR   18-APR-1997;   97GB-0007844.
PR   18-MAY-1996;   96GB-0010484.
XX
PA   (OXFO-) OXFORD VACS LTD.
XX
PI   Nuttall PA,  Paesen GC;
XX
DR   WPI; 1998-018506/02.
XX
PT   New vasoactive amine binding proteins and related nucleic acid,
PT   vectors - transformed cells and transgenic animals, used for
PT   assaying or removing histamine and as antihistamine or
PT   anti-inflammatory agents
XX
PS   Example 3; Page 20; 44pp; English.
XX
CC   This reverse primer was used with a forward primer (see AAV00233)
CC   to amplify the coding region (see AAV00227) of Rhipicephalus
CC   appendiculatus female-specific histamine binding protein 1
CC   (FS-HBP1) (see AAW37446), a novel vasoactive amine binding protein
CC   (VABP).  The primers were designed so that a SacI site was added
CC   upstream of the start codon, while the stop codon was replaced by
CC   a BamHI site, followed by 6 histidine codons and an SpeI site
CC   comprising a TAG stop codon.  The PCR product was ligated into
CC   transfer vector pAcC1291, generating plasmid pACC129.1-FS1.HIS.
CC   FS-HBP1 was expressed as a histidine-tagged protein in Spodoptera
CC   frugiperda Sf21 ovarian cells using a baculovirus expression system.
CC   VABPs can be used to assay or remove histamine, as an antihistamine
CC   or anti-inflammatory agent, and in vaccines.
XX
SQ   Sequence 50 BP; 11 A; 8 C; 14 G; 17 T; 0 other;


alignment_scores:
            Quality:   41.00              Length:     10
              Ratio:  4.100                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 70.000

alignment_block:
 US-09-439-311-2 x AAV00234   ..

 Align seg 1/1  to: AAV00234  from: 1  to: 50

     262 SerAspGlyAspGluAsnGlySerLeuIle 271
         ||||||||||||:::::|||||||||:::
       7 AGTGATGGTGATGATGATGGATCCCTTCTG 36

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ96924

seq_documentation_block:
ID   AAZ96924 standard; DNA; 59 BP.
XX
AC   AAZ96924;
XX
DT   14-APR-2000  (first entry)
XX
DE   S. cerevisiae gene deletion cassette constructing primer YMR290c-S1.
XX
KW   Antimycotic; mycosis; immunodepression; AIDS; diabetes; fungicide;
KW   mycete; gene deletion; PCR primer; ss.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO9955907-A2.
XX
PD   04-NOV-1999.
XX
PF   22-APR-1999;   99WO-EP02722.
XX
PR   24-APR-1998;   98EP-0401007.
PR   11-SEP-1998;   98EP-0402254.
XX
PA   (HMRI ) HOECHST MARION ROUSSEL.
XX
PI   Diu-Hercend A,  Entian K,  Koetter P;
XX
DR   WPI; 2000-105527/09.
XX
PT   Identifying antimycotic substances useful for drug preparation and
PT   treatment of mycosis -
XX
PS   Examples; Page 71; 86pp; English.
XX
CC   The invention provides a method of screening for antimycotic substances
CC   using essential genes from mycetes or a functionally similar mycete
CC   gene or the corresponding encoded protein as target. The essential gene
CC   useful for screening antimycotic substances is selected from the
CC   following genes: YML114c, YLR186w, YLR215c, YLR222c, YLR243w, YLR272c,
CC   YLR275w, YLR276c, YLR317w, YLR359w, YLR373c, YLR424w, YLR437c, YLR440c,
CC   YML023c, YML049c, YML077w, YML093w, YML127w, YMR032w, YMR093w, YMR131c,
CC   YMR185w, YMR212c, YMR213w, YMR218c, YMR281w, YMR288w, YMR290c, YMR211w,
CC   YMR049c, YMR134w, YDR196c, YDR299w, YDR365c, YDR396w, YDR407c, YDR416w,
CC   YDR449c, YDR472w, YDR499w, YDR141c, YDR324c, YDR325w, YDR398w, YDR246w,
```

```
CC   YDR236c, YDR361c, YDR367w, YDR339c, YDR413c, YDR429c, YDR468c, YDR489w,
CC   YDR527w, YDR288w, YDR201w, YDR434w, YDR181c, YDR531w, YPL126w, YPL093w,
CC   YPL063w, YPL024w, YPL020c, YPL012w, YPL007c, YPL233w, YPL146c, YIL091c,
CC   YIL083c, YIL019w, YIL109c, YIL104c, YFL024c, YFR003c, YFR027w, YFR042w,
CC   YIR010w, YIR015w, YPR048w, YPR072w, YPR082c, YPR085c, YPR105c, YPR112c,
CC   YPR137w, YPR143w, YPR144c and YPR169w. The method is useful for
CC   identifying substances for the preparation of drugs for the treatment of
CC   mycosis or prevention in immunodepression states. Drugs containing
CC   antimycotic substances are useful for the treatment of mycotic
CC   infections which occur during diseases like AIDS or diabetes. Substances
CC   which may be used for the fabrication of fungicides, especially of
CC   fungicides which are harmless for humans and animals and antimycotic
CC   substances which selectively inhibit the growth of specific mycete
CC   species only, can also be identified by this method. Sequences
CC   AAZ96811-Z96990 represent PCR primers used in construction of S.
CC   cerevisiae deletion cassettes.
XX
SQ   Sequence 59 BP; 8 A; 12 C; 15 G; 24 T; 0 other;


alignment_scores:
           Quality:   40.00            Length:     15
             Ratio:  3.077              Gaps:      0
 Percent Similarity:  86.667   Percent Identity: 53.333

alignment_block:
 US-09-439-311-2 x AAZ96924    ..

 Align seg 1/1  to: AAZ96924  from: 1  to: 59

     163 SerLysIleGlyValThrArgPheGluThrGlySerGlnSerPhe 177
         ::::::::::|||:::   ||||||:::|||||||||   ||||||
       9 ACGTCTTTTGGTATTTGGCGTTTTTCCACTGGCAGCTGAAGCTTC 53

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:AAT00254

seq_documentation_block:
ID   AAT00254 standard; DNA; 60 BP.
XX
AC   AAT00254;
XX
DT   14-AUG-1996  (first entry)
XX
DE   Thrombin 60N DNA ligand, clone #31.
XX
KW   Family 1; family 2; ligand; thrombin;
KW   systematic evolution of ligands by exponential enrichment; SELEX;
KW   heparin; selection; region of homology; inhibitor; ss.
XX
OS   Synthetic.
XX
PN   WO9521853-A1.
XX
PD   17-AUG-1995.
XX
PF   06-FEB-1995;   95WO-US01458.
XX
PR   28-MAR-1994;   94US-0219012.
PR   10-FEB-1994;   94US-0195005.
PR   11-JUN-1990;   90US-0536428.
PR   10-JUN-1991;   91US-0714131.
PR   22-APR-1993;   93US-0061691.
XX
PA   (NEXS-) NEXSTAR PHARM INC.
XX
PI   Gold L,  Janjic N,  Tasset D;
XX
DR   WPI; 1995-293073/38.
XX
PT   Identification of ligands to basic fibroblast growth factor and
PT   thrombin - which can be modified for increased in vivo stability
XX
PS   Claim 39; Page 97; 236pp; English.
XX
CC   The sequences given in AAT00202-25 and AAT00227-57 represent two groups
CC   of ligands to thrombin.  These sequences were isolated using the single
CC   stranded DNA molecules given in AAT00201 and AAT00226 which comprise a
CC   30N and a 60N variable region, respectively.  These ligands were
CC   isolated using systematic evolution of ligands by exponential enrichment
CC   (SELEX). The selection was conducted in a buffer solution at 37 deg. C.
CC   After 12 rounds of selection, no additional improvement in binding was
CC   seen.  By studying regions of homology between the isolated ligands, a
CC   truncated ligand of 38 nucleotides (see AAQ98403-04) was identified which
CC   retains high affinity binding and inhibits clotting.  These ligands are
CC   inhibitors of thrombin and are therefore useful in treating thrombin
CC   mediated conditions and in studying the structure and binding of
CC   thrombin.
XX
SQ   Sequence 60 BP; 10 A; 11 C; 29 G; 10 T; 0 other;


alignment_scores:
           Quality:   40.00            Length:     19
             Ratio:  2.667              Gaps:      0
 Percent Similarity:  78.947   Percent Identity: 42.105

alignment_block:
 US-09-439-311-2 x AAT00254    ..

 Align seg 1/1  to: AAT00254  from: 1  to: 60

     204 ThrSerValGlyThrGlyLeuGlyAlaLeuAlaGluGluIleAsnArgAs 220
         |||:::   |||:::||||||   ||||||||||:::   ::::::   ::
       4 ACCGCGGAGGGCGTAGGGTTGGAGGCGTTGGCCGATGTGGTAGGCACGGA 53

     220 nAlaAsp 222
         ::::|||
      54 CTCGGAT 60

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF70806

seq_documentation_block:
ID   AAF70806 standard; DNA; 60 BP.
XX
AC   AAF70806;
XX
DT   20-APR-2001  (first entry)
XX
DE   Thrombin high affinity ligand #53.
XX
KW   Ligand; basic fibroblast growth factor; bFGF; gene therapy; vascular;
KW   atherosclerosis; angioplasty; stability; ss.
XX
OS   Unidentified.
XX
PN   US6177557-B1.
XX
PD   23-JAN-2001.
XX
PF   05-AUG-1996;   96US-0687421.
XX
PR   11-JUN-1990;   90US-0536428.
PR   10-JUN-1991;   91US-0714131.
PR   06-NOV-1992;   92US-0973333.
PR   10-FEB-1994;   94US-0195005.
PR   28-MAR-1994;   94US-0219012.
XX
PA   (NEXS-) NEXSTAR PHARM INC.
XX
PI   Janjic N,  Gold L,  Tasset D;
XX
DR   WPI; 2001-158583/16.
XX
PT   Novel nucleic acid ligands to basic fibroblast growth factor that are
PT   useful as inhibitors of basic fibroblast growth factors and 2'-amino
PT   modified RNA ligands, exhibit increased in vivo stability  -
```

```
XX
PS   Example 19; Column 59-60; 153pp; English.
XX
CC   The present invention relates to a purified and isolated non-naturally
CC   occurring DNA ligands to basic fibroblast growth factor (bFGF).
CC   The ligands are useful as part of gene therapy treatments and
CC   for diagnosing pathogenesis of vascular diseases including
CC   initiation and progression of atherosclerosis, acute coronary
CC   syndromes, vein graft disease and restenosis following coronary
CC   angioplasty. The ligands have improved stability in vivo.
XX
SQ   Sequence 60 BP; 10 A; 11 C; 29 G; 10 T; 0 other;


alignment_scores:
            Quality:   40.00              Length:     19
              Ratio:  2.667                Gaps:      0
 Percent Similarity:  78.947   Percent Identity: 42.105

alignment_block:
 US-09-439-311-2 x AAF70806   ..

 Align seg 1/1  to: AAF70806  from: 1  to: 60

     204 ThrSerValGlyThrGlyLeuGlyAlaLeuAlaGluGluIleAsnArgAs 220
         |||:::   |||:::||||||   ||||||||||:::   ::::::   ::
       4 ACCGCGGAGGGCGTAGGGTTGGAGGCGTTGGCCGATGTGGTAGGCACGGA 53

     220 nAlaAsp 222
         ::::|||
      54 CTCGGAT 60

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI30690

seq_documentation_block:
ID   AAI30690 standard; DNA; 31 BP.
XX
AC   AAI30690;
XX
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) ATM 2.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,G)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 101; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 12 A; 4 C; 7 G; 8 T; 0 other;


alignment_scores:
            Quality:   39.00              Length:      9
              Ratio:  4.333                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 77.778

alignment_block:
 US-09-439-311-2 x AAI30690   ..

 Align seg 1/1  to: AAI30690  from: 1  to: 31

     139 ThrAsnGlnGluPheGlnIleGlySer 147
         ||||||:::||||||:::|||||||||
       5 ACAAATGAGGAATTCAGAATTGGTTCC 31

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV56429

seq_documentation_block:
ID   AAV56429 standard; DNA; 47 BP.
XX
AC   AAV56429;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Human ICAM-R cDNA primer #27.
XX
KW   Intercellular adhesion molecule; ICAM-R; human; modulator; 14.3.3 family;
KW   HS1-beta; tubulin; inhibitor; stimulator; effector; immune response;
KW   inflammation; disorder; T cell activation; macrophage; Crohn's disease;
KW   adult respiratory distress syndrome; stroke; multiple sclerosis; asthma;
KW   rheumatoid arthritis; tumour growth; human immune deficiency virus;
KW   infection; diabetes; graft vs. host disease; passive immunisation;
KW   primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   US5773218-A.
XX
PD   30-JUN-1998.
XX
PF   07-JUN-1995;   95US-0482882.
XX
PR   05-AUG-1994;   94US-0286754.
PR   27-JAN-1992;   92US-0827689.
PR   26-MAY-1992;   92US-0889724.
PR   05-JUN-1992;   92US-0894061.
PR   22-JAN-1993;   93US-0009266.
PR   26-JAN-1993;   93WO-US00787.
PR   05-AUG-1993;   93US-0102852.
PR   07-JUN-1995;   95US-0482882.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Vazeux R;
XX
DR   WPI; 1998-386989/33.
XX
PT   Identifying compounds that modulate interaction of intercellular
```

```
PT   adhesion molecule R - with ligands HS1-beta and tubulin using
PT   two-hybrid assay, useful for treating inflammation, T cell
PT   activation etc.
XX
PS   Example 13; Column 135-136; 108pp; English.
XX
CC   AAV56429-V56434 are primers used in the isolation of a novel human
CC   intercellular adhesion molecule, ICAM-R. This sequence is used in a
CC   method which investigates modulators of the interaction between ICAM-R
CC   and the 14.3.3 family member HS1-beta and tubulin. An anti-ICAM-R
CC   antibody optionally coupled to toxin or radionuclide, or an ICAM-R
CC   peptide, can block, inhibit or stimulate ligand/receptor interactions
CC   involving ICAM-R, particularly its effector functions involved in
CC   (non)specific immune responses. ICAM-R related agents may be used to
CC   treat or monitor inflammation, disorders involving T cell activation or
CC   macrophages, e.g. adult respiratory distress syndrome, stroke, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis, asthma, tumour
CC   growth, human immune deficiency virus infection, diabetes, graft vs. host
CC   disease and many others. Antibodies may also be used for passive
CC   immunisation, for purifying, detecting or quantifying ICAM-R and for
CC   identifying ICAM-R expressing cells.
XX
SQ   Sequence 47 BP; 9 A; 21 C; 7 G; 10 T; 0 other;


alignment_scores:
            Quality:   39.00            Length:     12
              Ratio:  3.900              Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AAV56429/rev   ..

 Align seg 1/1  to reverse of: AAV56429  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX21884

seq_documentation_block:
ID   AAX21884 standard; DNA; 47 BP.
XX
AC   AAX21884;
XX
DT   14-MAY-1999  (first entry)
XX
DE   Primer for antibody against ICAM-R.
XX
KW   ICAM; immunoglobulin-like loop; intercellular adhesion molecule receptor;
KW   alpha d/CD18; antibody; immunisation; inflammatory response; asthma;
KW   tumour growth; viral infection; therapy; primer; ss.
XX
OS   Synthetic.
OS   Mus sp.
XX
PN   US5880268-A.
XX
PD   09-MAR-1999.
XX
PF   07-JUN-1995;   95US-0483932.
XX
PR   05-AUG-1994;   94US-0286754.
PR   27-JAN-1992;   92US-0827689.
PR   26-MAY-1992;   92US-0889724.
PR   05-JUN-1992;   92US-0894061.
PR   22-JAN-1993;   93US-0009266.
PR   26-JAN-1993;   93WO-US00787.
PR   05-AUG-1993;   93US-0102852.
PR   07-JUN-1995;   95US-0483932.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Vazeux R;
XX
DR   WPI; 1999-204041/17.
XX
PT   New intercellular adhesion molecule receptor (ICAM-R) specific
PT   antibodies - useful for modulating ligand/receptor binding and
PT   biological activities involving ICAM-R, especially those of the
PT   specific and non-specific immune systems
XX
PS   Example 13; Column 41; 108pp; English.
XX
CC   This sequence is a primer for DNA encoding an antibody specific for
CC   ICAM-R. The invention relates to antibodies (Ab) which bind specifically
CC   to the intercellular adhesion molecule receptor (ICAM-R), inhibiting the
CC   interaction between ICAM-R and alpha d/CD18. Abs with specific ICAM-R
CC   binding are useful in compositions for immunisation, and for purifying
CC   ICAM-R polypeptides and identifying cells expressing ICAM-R on their cell
CC   surface, modulating ligand/receptor binding and biological activities
CC   involving ICAM-R, especially inflammatory responses of the specific
CC   immune system, the non-specific immune system, monitoring and treating
CC   asthma, tumour growth, and/or metastasis, and viral infection (e.g. HIV
CC   infection). In particular diseases involving an essential T cell
CC   activation (e.g. asthma, psoriasis, diabetes, graft vs. host disease,
CC   tissue transplant rejection, and multiple sclerosis) may be treated with
CC   anti-ICAM-R antibodies. The Abs specifically bind to and identify ICAM-R
CC   and disrupt ICAM-R to cell adhesion molecule, especially alpha d/CD18
CC   binding.
XX
SQ   Sequence 47 BP; 9 A; 21 C; 7 G; 10 T; 0 other;


alignment_scores:
            Quality:   39.00            Length:     12
              Ratio:  3.900              Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AAX21884/rev   ..

 Align seg 1/1  to reverse of: AAX21884  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV69197

seq_documentation_block:
ID   AAV69197 standard; DNA; 47 BP.
XX
AC   AAV69197;
XX
DT   17-FEB-1999  (first entry)
XX
DE   Humanised ICR-1.1 antibody Vk region DNA mutating oligo 110.
XX
KW   Intercellular adhesion molecule polypeptide; ICAM-R; humanised; ICR 1.1;
KW   ICR 8.1; monoclonal antibody; therapeutic; inflammatory; asthma; tumour;
KW   graft-versus-host disease; viral infection; toxin; radionuclide;
KW   neovascularisation site; mutagenic; PCR primer; ss.
XX
OS   Synthetic.
OS   Mus sp.
XX
PN   US5837822-A.
XX
PD   17-NOV-1998.
XX
PF   07-JUN-1995;   95US-0487113.
XX
PR   07-JUN-1995;   95US-0487113.
PR   27-JAN-1992;   92US-0827689.
```

```
PR   26-MAY-1992;   92US-0889724.
PR   05-JUN-1992;   92US-0894061.
PR   22-JAN-1993;   93US-0009266.
PR   26-JAN-1993;   93WO-US00787.
PR   05-AUG-1993;   93US-0102852.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Vazeux R;
XX
DR   WPI; 1999-023535/02.
XX
PT   Humanised antibodies specific for intercellular adhesion molecule
PT   polypeptide - useful for therapeutic or diagnostic purposes
XX
PS   Example 13; Column 42; 116pp; English.
XX
CC   The invention relates to humanised ICR 1.1 and ICR 8.1 antibodies
CC   targeted to the human intercellular adhesion molecule polypeptide
CC   (ICAM-R) polypeptide. Antibodies specific for ICAM's are potentially
CC   useful as therapeutic compounds, for treating e.g. immune-mediated
CC   inflammatory conditions (e.g. graft-versus-host disease), asthma,
CC   tumours or viral infections. Monoclonal antibodies specific for ICAM-R,
CC   or their conjugates formed with e.g. toxins or radionuclides are useful
CC   for therapeutically targeting or detecting neovascularisation sites.
CC   PCR mutagenic oligos AAV69197 and AAV69198 are used in the construction
CC   of the Vk region of the humanised antibody ICR-1.1.
XX
SQ   Sequence 47 BP; 9 A; 21 C; 7 G; 10 T; 0 other;


alignment_scores:
            Quality:   39.00              Length:     12
              Ratio:  3.900                Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 66.667

alignment_block:
 US-09-439-311-2 x AAV69197/rev   ..

 Align seg 1/1  to reverse of: AAV69197  from: 1  to: 47

     169 ArgPheGluThrGlySerGlnSerPheSerSerGly 180
         |||:::|||||||||||||      |||:::||||||
      44 AGGATGGAGACTGGGTCAGCACGATTTGGGAGTGGA 9
```

```
OM of: US-09-439-311-2 to: EST:*   out_format : pfs

Date: Apr 17, 2002  2:43 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09439311/runat_16042002_134010_11677/app_query.fasta_1.395
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=60
 -USER=US09439311_@CGN1_1_3691 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-439-311-2
Query length: 333
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 1467.110000



score_list:
Sequence              Strd Orig    ZScore    EScore Len   ! Documentation  ..
gb_est1:AU104260       +   48.00   103.91  1.3e+04   50   ! AU104260 AU104260 Sugano Homo sa
gb_gss:B04096          +   45.00    96.72  3.2e+04   60   ! B04096 cSRL-25g1-u cSRL flow sor
gb_gss:AZ469793        +   39.00    88.62  9.0e+04   44   ! AZ469793 1M0283F04R Mouse 10kb p
gb_est1:AU107968       +   39.00    87.43  1.0e+05   50   ! AU107968 AU107968 Sugano Homo sa
gb_est2:W20078         +   39.00    87.07  1.1e+05   52   ! W20078 zb40e03.r1 Soares_parathy
gb_gss:AZ921603        -   39.00    86.21  1.2e+05   57   ! AZ921603 1006030F10.x1 1006 - Re
gb_gss:AZ998589        +   38.50    85.80  1.3e+05   54   ! AZ998589 2M0285D08R Mouse 10kb p
gb_est1:AU106648       -   38.00    85.60  1.3e+05   50   ! AU106648 AU106648 Sugano Homo sa
gb_est2:C20861         -   38.00    84.55  1.5e+05   56   ! C20861 HUMGS0004926 Human adult
gb_gss:AZ654882        -   38.00    84.38  1.5e+05   57   ! AZ654882 1M0529N22F Mouse 10kb p
gb_est1:AA628048       +   38.00    83.90  1.6e+05   60   ! AA628048 nq62f05.s1 NCI_CGAP_Ov
gb_gss:AZ429658        +   37.00    87.63  1.0e+05   33   ! AZ429658 1M0213A18R Mouse 10kb p
gb_est1:AI595782       +   37.00    83.96  1.6e+05   49   ! AI595782 uj69c12.x1 Sugano mouse
gb_est1:AU102738       -   37.00    83.77  1.7e+05   50   ! AU102738 AU102738 Sugano Homo sa
gb_est1:AU102739       -   37.00    83.77  1.7e+05   50   ! AU102739 AU102739 Sugano Homo sa
gb_gss:CNS040OQ        +   37.00    83.23  1.8e+05   53   ! AL269171 Tetraodon nigroviridis
gb_est1:AA664752       -   37.00    82.55  2.0e+05   57   ! AA664752 nu65c03.s1 NCI_CGAP_Alv
gb_est1:AI320166       +   37.00    82.39  2.0e+05   58   ! AI320166 c1e11nm.f1 Neurospora c
gb_gss:AZ996603        -   36.50    81.47  2.3e+05   58   ! AZ996603 2M0282P18R Mouse 10kb p
gb_gss:AZ615427        +   36.00    83.78  1.7e+05   41   ! AZ615427 1M0444G16R Mouse 10kb p
gb_est1:AI522735       +   36.00    82.71  1.9e+05   46   ! AI522735 fb61e01.x1 Zebrafish Wa
gb_gss:AZ328471        -   36.00    82.32  2.0e+05   48   ! AZ328471 1M0052F16F Mouse 10kb p
gb_est1:AU105900       -   36.00    81.94  2.1e+05   50   ! AU105900 AU105900 Sugano Homo sa
gb_est1:AI905591       +   36.00    81.75  2.2e+05   51   ! AI905591 CM-BT094-050299-117 BT0
gb_gss:AZ402345        -   36.00    81.75  2.2e+05   51   ! AZ402345 1M0169B10R Mouse 10kb p
gb_est1:BE043282       +   36.00    81.57  2.2e+05   52   ! BE043282 hk48h12.y1 NCI_CGAP_Ov3
gb_est1:AA673303       +   36.00    81.05  2.4e+05   55   ! AA673303 vn69f08.r1 Barstead mou
gb_est1:AA711965       +   36.00    81.05  2.4e+05   55   ! AA711965 vu29c05.r1 Barstead mou
gb_est1:AI444123       +   36.00    81.05  2.4e+05   55   ! AI444123 fb26b06.x1 Zebrafish Wa
gb_est1:AI831322       +   36.00    81.05  2.4e+05   55   ! AI831322 wj82d10.x1 NCI_CGAP_Lym
gb_gss:AZ412625        -   35.50    79.48  2.9e+05   59   ! AZ412625 1M0186L07F Mouse 10kb p
gb_est1:AI633469       +   35.00    81.51  2.2e+05   43   ! AI633469 th62d06.x1 NCI_CGAP_Ov2
gb_est2:W53880         -   35.00    81.51  2.2e+05   43   ! W53880 md03h05.r1 Soares mouse e
gb_est1:AU007710       +   35.00    80.29  2.6e+05   49   ! AU007710 AU007710 Schizosaccharo
gb_est1:AA400193       -   35.00    80.29  2.6e+05   49   ! AA400193 zu64e08.s1 Soares_testi
gb_est1:AU106926       -   35.00    80.11  2.7e+05   50   ! AU106926 AU106926 Sugano Homo sa
gb_est1:AI653077       +   35.00    79.74  2.8e+05   52   ! AI653077 wb23c01.x1 NCI_CGAP_GC6
gb_est1:AI889129       -   35.00    79.74  2.8e+05   52   ! AI889129 tw36h11.x1 NCI_CGAP_Ut1
gb_est1:AA449400       +   35.00    79.22  3.0e+05   55   ! AA449400 zx04f01.r1 Soares_total
gb_est2:BG654043       -   35.00    78.89  3.1e+05   57   ! BG654043 sad50f03.y2 Gm-c1075 Gl
gb_est2:BG099588       -   35.00    78.89  3.1e+05   57   ! BG099588 nag54e01.x1 NCI_CGAP_Co
gb_gss:AZ372199        +   35.00    78.41  3.3e+05   60   ! AZ372199 1M0123N23R Mouse 10
gb_est1:AA620370       +   34.50    79.38  2.9e+05   49   ! AA620370 af07f05.s1 Soares_t
gb_est1:AW697131       +   34.50    78.83  3.1e+05   52   ! AW697131 NF115E02ST1F1017 De
gb_gss:AZ309933        -   34.00    78.46  3.3e+05   49   ! AZ309933 1M0017G23F Mouse 10

seq_name: gb_est1:AU104260

seq_documentation_block:
LOCUS       AU104260        50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU104260 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP15388, mRNA sequence.
ACCESSION   AU104260
VERSION     AU104260.1  GI:13553781
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP15388"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       17 a      12 c      11 g      10 t
ORIGIN


alignment_scores:
           Quality:   48.00            Length:     14
             Ratio:  4.364              Gaps:      0
 Percent Similarity:  78.571   Percent Identity: 71.429

alignment_block:
 US-09-439-311-2 x AU104260   ..

 Align seg 1/1  to: AU104260  from: 1  to: 50

     143 PheGlnIleGlySerSerSerAsnGlnThrIleLysAlaSer 156
         ||||||:::|||      ||||||||||||   |||||||||
       4 TTTCAGCTTGGACACTGTTCCAATCAGACAAGCAAAGCAAGC 45

seq_name: gb_gss:B04096

seq_documentation_block:
LOCUS       B04096          60 bp    DNA             GSS       13-JUL-1996
DEFINITION  cSRL-25g1-u cSRL flow sorted Chromosome 11 specific cosmid Homo
            sapiens genomic clone cSRL-25g1, DNA sequence.
ACCESSION   B04096
VERSION     B04096.1  GI:1413374
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,
            Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S. ., Harris
```

```
          ,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and
          Garner,H.R.
  TITLE   Genomic Sequence Sampled Map of Chromosome 11
  JOURNAL Unpublished (1996)
COMMENT   Contact: Evans GA, Shane Probst
          McDermott Center for Human Growth and Development
          University of Texas Southwestern Medical Center At Dallas
          5323 Harry Hines Blvd, Dallas TX 75235-8591
          Tel: 214-648-1600
          Fax: 214-648-1666
          Email: gevans@utsw.swmed.edu, shane@mcdermott.swmed.edu
          Seq primer: T7
          Class: cosmid ends
          High quality sequence stop: 60.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="cSRL-25g1"
                     /clone_lib="cSRL flow sorted Chromosome 11 specific
                     cosmid"
                     /sex="female"
                     /cell_type="chimeric hamster somatic cell hybrid"
                     /note="Vector: sCos-1; Human Chromosome 11 specific cosmid
                     library prepared from flow sorted human  Chromosome 11
                     derived from Chinese Hampster Ovary (CHO) monochromosomal
                     somatic cell hybrid, J1"
BASE COUNT       16 a     16 c      5 g     22 t      1 others
ORIGIN


alignment_scores:
           Quality:   45.00             Length:     17
             Ratio:  3.000               Gaps:      0
 Percent Similarity:  88.235   Percent Identity: 52.941

alignment_block:
 US-09-439-311-2 x B04096   ..

 Align seg 1/1  to: B04096  from: 1  to: 60

     112 IleAsnArgLeuMetGluGluLeuAspAsnIleAlaAsnThrThrSerPh 128
         |||:::   |||:::|||:::|||   :::|||:::||||||:::|||||
       4 ATTACAGCTCTCATTGAGGACCTCCCTTCTATTTCNAATACAAATAGTTT 53

     128 e 128
         |
      54 C 54

seq_name: gb_gss:AZ469793

seq_documentation_block:
LOCUS       AZ469793       44 bp    DNA             GSS       04-OCT-2000
DEFINITION  1M0283F04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0283F04 R, DNA sequence.
ACCESSION   AZ469793
VERSION     AZ469793.1  GI:10627918
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0283  row: F  column: 04
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 44.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0283F04"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        9 a     12 c     12 g     11 t
ORIGIN


alignment_scores:
           Quality:   39.00             Length:     13
             Ratio:  3.545               Gaps:      0
 Percent Similarity:  84.615   Percent Identity: 53.846

alignment_block:
 US-09-439-311-2 x AZ469793   ..

 Align seg 1/1  to: AZ469793  from: 1  to: 44

     312 GlyAlaGlyIleLeuHisThrGluAsnTyrGlyArgLeu 324
         ||||||||||||       :::::::|||:::|||:::|||
       5 GGGGCTGGCATTGCTCTCAATGACAACTTTGGCAAGCTC 43

seq_name: gb_est1:AU107968

seq_documentation_block:
LOCUS       AU107968       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU107968 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            KAT11118, mRNA sequence.
ACCESSION   AU107968
VERSION     AU107968.1  GI:13557490
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
```

```
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="KAT11118"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       10 a     11 c     15 g     14 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     13
              Ratio:  3.900               Gaps:      0
 Percent Similarity:  76.923   Percent Identity: 61.538

alignment_block:
 US-09-439-311-2 x AU107968    ..

 Align seg 1/1  to: AU107968  from: 1  to: 50

     305 IleThrGlySerIleGlyValGlyAlaGlyIleLeuHis 317
         ||||||:::      |||||||||:::|||   ||||||
       3 ATTACCGCTTGTTCTGGAGTCGGAGTTGGAACGCTCCAC 41

seq_name: gb_est2:W20078

seq_documentation_block:
LOCUS       W20078         52 bp    mRNA            EST       03-MAY-1996
DEFINITION  zb40e03.r1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:306076 5' similar to SW:ATPL_BOVIN P32876 ATP SYNTHASE
            LIPID-BINDING PROTEIN P1 PRECURSOR ;, mRNA sequence.
ACCESSION   W20078
VERSION     W20078.1  GI:1295949
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: mob.REGA+ET
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Homo sapiens"
                     /db_xref="GDB:1249486"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:306076"
                     /clone_lib="Soares_parathyroid_tumor_NbHPA"
                     /tissue_type="parathyroid tumor"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTT
                     TTTTT-3'], double-stranded cDNA was size selected, ligated
                     to Eco RI adapters (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of a modified pT7T3
                     vector (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                     adenomas was kindly provided by Dr. Stephen Marx, National
                     Institute of Diabetes and Digestive and Kidney Diseases,
                     NIH."
BASE COUNT       8 a      7 c     19 g     14 t      4 others
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     17
              Ratio:  3.545               Gaps:      0
 Percent Similarity:  64.706   Percent Identity: 47.059

alignment_block:
 US-09-439-311-2 x W20078    ..

 Align seg 1/1  to: W20078  from: 1  to: 52

     299 AspGlyArgGlyIleLysIleThrGlySerIleGlyValGlyAlaGlyIl 315
         :::|||   |||         :::|||:::   |||   ||||||||||
       2 AATGGTGNAGGTGCTNCAACAGTAGGCNTGGCTGGTNCTGGTGCTGGGAT 51

     315 e 315
         |
      52 T 52

seq_name: gb_gss:AZ921603

seq_documentation_block:
LOCUS       AZ921603       57 bp    DNA             GSS       20-MAR-2001
DEFINITION  1006030F10.x1 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   AZ921603
VERSION     AZ921603.1  GI:13393406
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Sequence was trimmed at possible ligation site. Post-ligation
            sequence submitted separately.
            Plate: 1006030  row: F  column: 10
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Zea mays"
```

```
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       13 a     18 c      7 g     19 t
ORIGIN


alignment_scores:
            Quality:   39.00             Length:     11
              Ratio:  4.333               Gaps:      0
 Percent Similarity:  81.818   Percent Identity: 54.545

alignment_block:
 US-09-439-311-2 x AZ921603/rev   ..

 Align seg 1/1  to reverse of: AZ921603  from: 1  to: 57

    309 IleGlyValGlyAlaGlyIleLeuHisThrGlu 319
        :::|||:::|||:::|||   ||||||   |||
     54 CTGGGGATAGGAAGTGGATATCTGCACAAGGAG 22

seq_name: gb_gss:AZ998589

seq_documentation_block:
LOCUS        AZ998589       54 bp    DNA             GSS       27-APR-2001
DEFINITION   2M0285D08R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
             clone UUGC2M0285D08 R, DNA sequence.
ACCESSION    AZ998589
VERSION      AZ998589.1  GI:13869816
KEYWORDS     GSS.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 54)
  AUTHORS    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
             Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
             ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
             and Wright,D.,Weiss,R.
  TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
             plasmid inserts
  JOURNAL    Unpublished (2000)
COMMENT      Contact: Robert B. Weiss
             University of Utah Genome Center
             University of Utah
             Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
             84112, USA
             Tel: 801 585 5606
             Fax: 801 585 7177
             Email: ddunn@genetics.utah.edu
             Insert Length: 10000   Std Error: 0.00
             Plate: 0285  row: D  column: 08
             Seq primer: CACACAGGAAACAGCTATGACC
             Class: plasmid ends
             High quality sequence stop: 54.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0285D08"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       11 a      2 c     21 g     20 t
ORIGIN


alignment_scores:
            Quality:   38.50             Length:     18
              Ratio:  2.750               Gaps:      1
 Percent Similarity:  77.778   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x AZ998589   ..

 Align seg 1/1  to: AZ998589  from: 1  to: 54

    194 AspPheLysPheAspSerValValIleSerThrSerValGlyThrGlyLe 210
        |||:::||||||:::|||:::|||:::         ||||||   |||::
     10 GATTATAAATTTGAATCTCTTGTATTT.........GTGGGGTAGGGGGT 50

    210 uGly 211
        :|||
     51 GGGG 54

seq_name: gb_est1:AU106648

seq_documentation_block:
LOCUS        AU106648       50 bp    mRNA            EST       05-APR-2001
DEFINITION   AU106648 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
             KAT05523, mRNA sequence.
ACCESSION    AU106648
VERSION      AU106648.1  GI:13556169
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 50)
  AUTHORS    Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
             ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
             ,K., Suyama,A. and Sugano,S.
  TITLE      Fine Structural analysis of transcription start sites of human
             mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL    Unpublished (2001)
COMMENT      Contact: Yutaka Suzuki
             Department of Virology
             Institute of Medical Science, University of Tokyo
             4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
             Email: ysuzuki@ims.u-tokyo.ac.jp
             Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
```

```
          ,S. Construction and characterization of a full length-enriched and
          a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="KAT05523"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       10 a      20 c       8 g      12 t
ORIGIN


alignment_scores:
            Quality:   38.00              Length:     16
              Ratio:  3.455                Gaps:      0
 Percent Similarity:  68.750   Percent Identity: 56.250

alignment_block:
 US-09-439-311-2 x AU106648/rev   ..

 Align seg 1/1  to reverse of: AU106648  from: 1  to: 50

     247 GlnAspPheAlaIleAsnGlyValValIleGlyLysValAspTyrSer 262
         |||      |||::::::|||   |||   |||   ||||||||||||
      49 CAGGCTGCAGCTGTGTCAGGTAGAGTGCAAGGGGGAGTAGACTATTCG 2

seq_name: gb_est2:C20861

seq_documentation_block:
LOCUS        C20861          56 bp    mRNA            EST       23-OCT-1996
DEFINITION   HUMGS0004926 Human adult (K.Okubo) Homo sapiens cDNA 3', mRNA
             sequence.
ACCESSION    C20861
VERSION      C20861.1  GI:1621971
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 56)
  AUTHORS    Okubo,K.
  TITLE      BodyMap; human gene expression database
  JOURNAL    Unpublished (1995)
COMMENT      Contact: Okubo,K.
             Institute for Molecular and Cellular Biol
             Osaka University
             1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan
             Tel: 06-877-5111(ex.3315)
             Email: kousaku@imcb.osaka-u.ac.jp
             Human Gene Signature, 3'-directed cDNA sequence. We are not
             submitting the same cDNA sequence redundantly to DDBJ since 1993.
             For the abundance information of clones with this sequence in this
             library and as well as in other 3'-directed libraries, see '
             http://www.imcb.osaka-u.ac.jp/bodymap'. The sequences of the clones
             represented by this GS sequences is also found there.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Human adult (K.Okubo)"
                     /dev_stage="adult"
BASE COUNT       19 a       9 c       7 g      19 t       2 others
ORIGIN


alignment_scores:
            Quality:   38.00              Length:     18
              Ratio:  2.923                Gaps:      0
 Percent Similarity:  72.222   Percent Identity: 44.444

alignment_block:
 US-09-439-311-2 x C20861/rev   ..

 Align seg 1/1  to reverse of: C20861  from: 1  to: 56

     215 GluGluIleAsnArgAsnAlaAspLysThrGlyIleArgAlaThrPheAs 231
         |||:::|||      |||:::     ||||||||||||   :::|||::
      54 GAACAGATATATTCTAATACAGTTGTCACTGGTATAAGATTTCNNTTTGA 5

     231 pVal 232
         ::::
       4 GATC 1

seq_name: gb_gss:AZ654882

seq_documentation_block:
LOCUS        AZ654882       57 bp    DNA             GSS       14-DEC-2000
DEFINITION   1M0529N22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
             clone UUGC1M0529N22 F, DNA sequence.
ACCESSION    AZ654882
VERSION      AZ654882.1  GI:11792028
KEYWORDS     GSS.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 57)
  AUTHORS    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
             Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
             ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
             and Wright,D.,Weiss,R.
  TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
             plasmid inserts
  JOURNAL    Unpublished (2000)
COMMENT      Contact: Robert B. Weiss
             University of Utah Genome Center
             University of Utah
             Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
             84112, USA
             Tel: 801 585 5606
             Fax: 801 585 7177
             Email: ddunn@genetics.utah.edu
             Insert Length: 10000   Std Error: 0.00
             Plate: 0529  row: N  column: 22
             Seq primer: CGTTGTAAAACGACGGCCAGT
             Class: plasmid ends
             High quality sequence stop: 57.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0529N22"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       16 a      23 c       6 g      12 t
```

```
ORIGIN


alignment_scores:
            Quality:   38.00              Length:     14
              Ratio:  3.800                Gaps:      0
 Percent Similarity:  71.429   Percent Identity: 42.857

alignment_block:
 US-09-439-311-2 x AZ654882/rev   ..

 Align seg 1/1  to reverse of: AZ654882  from: 1  to: 57

     302 GlyIleLysIleThrGlySerIleGlyValGlyAlaGlyIle 315
         |||   :::|||   |||   :::|||:::|||   |||:::
      43 GGCAGTAGGATATTGGGTGTGTTGGGCATAGGCCAAGGAGTG 2

seq_name: gb_est1:AA628048

seq_documentation_block:
LOCUS       AA628048       60 bp    mRNA            EST       31-OCT-1997
DEFINITION  nq62f05.s1 NCI_CGAP_Ov6 Homo sapiens cDNA clone IMAGE:1154625
            similar to gb:U07857 14 KD PROTEIN OF SIGNAL RECOGNITION PARTICLE
            (HUMAN);, mRNA sequence.
ACCESSION   AA628048
VERSION     AA628048.1  GI:2540047
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Andrew Berchuck, M.D., Elise Kohn, M.D.,
            Rodrigo F. Chuaqui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 300   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1154625"
                     /clone_lib="NCI_CGAP_Ov6"
                     /sex="female"
                     /tissue_type="normal cortical stroma"
                     /lab_host="DH10B"
                     /note="Organ: ovary; Vector: pAMP10; mRNA made from normal
                     ovarian cortical stroma, cDNA made by oligo-dT  priming.
                     Non-directionally cloned.  Size-selected on agarose  gel,
                     average insert size 600 bp.  "
BASE COUNT       18 a     14 c     12 g     16 t
ORIGIN


alignment_scores:
            Quality:   38.00              Length:     12
              Ratio:  3.455                Gaps:      0
 Percent Similarity:  91.667   Percent Identity: 58.333
```

```
alignment_block:
 US-09-439-311-2 x AA628048   ..

 Align seg 1/1  to: AA628048  from: 1  to: 60

     180 GlyThrValGlyLeuThrIleLysAsnTyrAsnGly 191
         |||:::|||   :::||||||||||:::|||:::|||
       7 GGCAGCGTCTATATCACCATTAAGAAGTATGACGGT 42

seq_name: gb_gss:AZ429658

seq_documentation_block:
LOCUS       AZ429658       33 bp    DNA             GSS       03-OCT-2000
DEFINITION  1M0213A18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0213A18 R, DNA sequence.
ACCESSION   AZ429658
VERSION     AZ429658.1  GI:10553671
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0213  row: A  column: 18
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 33.
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0213A18"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        9 a      5 c     11 g      8 t
ORIGIN
```

```
alignment_scores:
            Quality:   37.00              Length:     10
              Ratio:  4.111                Gaps:      0
 Percent Similarity:  90.000   Percent Identity: 70.000

alignment_block:
 US-09-439-311-2 x AZ429658   ..

 Align seg 1/1  to: AZ429658  from: 1  to: 33

     263 AspGlyAspGluAsnGlySerLeuIleSer 272
         ||||||   |||:::||||||:::||||||
       3 GATGGGGGAGAACATGGATCTATGATCTCC 32

seq_name: gb_est1:AI595782

seq_documentation_block:
LOCUS       AI595782       49 bp    mRNA            EST       21-APR-1999
DEFINITION  uj69c12.x1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1925206 3' similar to TR:P70464 P70464 ALPHA-2-GLYCOPROTEIN
            1, ZINC ;, mRNA sequence.
ACCESSION   AI595782
VERSION     AI595782.1  GI:4604830
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:981498
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1925206"
                     /clone_lib="Sugano mouse liver mlia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT       10 a      12 c      11 g      16 t
ORIGIN


alignment_scores:
            Quality:   37.00              Length:     12
              Ratio:  3.700                Gaps:      0
 Percent Similarity:  83.333   Percent Identity: 58.333

alignment_block:
 US-09-439-311-2 x AI595782   ..

 Align seg 1/1  to: AI595782  from: 1  to: 49

     126 ThrSerPheAsnGlyLysGlnLeuLeuSerGlyGly 137
         |||   |||:::|||::::::|||   |||||||||
       7 ACACCTTTTTCTGGTCAAATGCTAGCTTCTGGTGGC 42

seq_name: gb_est1:AU102738

seq_documentation_block:
LOCUS       AU102738       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU102738 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC13040, mRNA sequence.
ACCESSION   AU102738
VERSION     AU102738.1  GI:13552259
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC13040"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        7 a      16 c      18 g       9 t
ORIGIN


alignment_scores:
            Quality:   37.00              Length:     14
              Ratio:  2.846                Gaps:      0
 Percent Similarity:  92.857   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x AU102738/rev   ..

 Align seg 1/1  to reverse of: AU102738  from: 1  to: 50

      93 ThrGlnAlaAlaGlnAspGlyGlnSerLeuLysThrArgThr 106
         |||:::|||||||:::   |||:::|||::::::::::|||||
      42 ACTAGGGCCGCGCGTACAGGGAGCTCCATTTCCTCCCGCACG 1
```

```
seq_name: gb_est1:AU102739

seq_documentation_block:
LOCUS       AU102739       50 bp    mRNA            EST       05-APR-2001
DEFINITION  AU102739 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC13119, mRNA sequence.
ACCESSION   AU102739
VERSION     AU102739.1  GI:13552260
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
  TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC13119"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        7 a     16 c     18 g      9 t
ORIGIN


alignment_scores:
            Quality:   37.00            Length:     14
              Ratio:  2.846              Gaps:      0
 Percent Similarity:  92.857   Percent Identity: 50.000

alignment_block:
 US-09-439-311-2 x AU102739/rev   ..

 Align seg 1/1  to reverse of: AU102739  from: 1  to: 50

      93 ThrGlnAlaAlaGlnAspGlyGlnSerLeuLysThrArgThr 106
         |||:::||||||:::   |||:::|||:::::::::||||||
      42 ACTAGGGCCGCGCGTACAGGGAGCTCCATTTCCTCCCGCACG 1
```